

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE PSICOLOGÍA

**Personalidad, Evaluación y Tratamiento Psicológico I (Personalidad,
Evaluación y Psicología Clínica)**



TESIS DOCTORAL

**Validación de la versión española del BDI-FS: un instrumento para
la detección de la depresión en atención primaria**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Amaia Izquierdo Elizo

Directores

**Jesús Sanz Fernández
María Paz García Vera**

Madrid, 2011

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Psicología

Departamento de Personalidad, Evaluación y Tratamiento
Psicológico I (Personalidad, Evaluación y Psicología Clínica).

TESIS DOCTORAL



***Validación de la versión española del BDI - FS:
un instrumento para la detección de la depresión
en atención primaria.***

AUTORA:

Amaia Izquierdo Elizo

DIRECTORES:

Jesús Sanz Fernández

María Paz García Vera

Madrid, 2011

***Dedicada a mis padres, Paco y Piedad,
gracias por cultivar en mí la inquietud
por el conocimiento.***

ÍNDICE

<u>PARTE TEÓRICA.</u>	1
CAPITULO 1:TRASTORNOS DEPRESIVOS EN ATENCIÓN PRIMARIA:	3
1.1.- <u>Conceptualización y clasificación de los trastornos depresivos.</u>	4
1.1.1.-Conceptualización de los trastornos depresivos.	4
1.1.2.- Clasificación de los trastornos depresivos.	13
1.2.- <u>Prevalencia de la depresión en población general.</u>	17
1.2.1.-Prevalencia de los trastornos depresivos en atención primaria.	25
1.2.2.-Comorbilidad de los trastornos depresivos en atención primaria.	29
1.3.- <u>Importancia de los trastornos depresivos.</u>	33
1.3.1.-Repercusiones de los trastornos depresivos en la sociedad.	33
1.3.2.-Repercusiones a nivel de salud pública.	36
1.4.- <u>Dificultades en el diagnóstico de los trastornos depresivos en Atención Primaria.</u>	38
1.4.1.- Infradiagnóstico y sobrediagnóstico de los trastornos depresivos en atención primaria.	39
1.4.2.- Factores asociados a las dificultades diagnósticas de la depresión.	42
1.5.- <u>Métodos de detección de la depresión por parte de los médicos de atención primaria.</u>	46
 CAPÍTULO 2: USO DE INSTRUMENTOS DE CRIBADO O DESPISTAJE (SCREENING) DE LOS TRASTORNOS DEPRESIVOS.	 49
2.1.- <u>Conceptualización y uso de los instrumentos de cribado o despistaje (screening).</u>	50

2.1.1.- Definición de cribado o despistaje (<i>screening</i>).	52
2.1.2.- Usos habituales del cribado.	57
2.2.- <u>Población diana susceptible del cribado de los trastornos depresivos.</u>	60
2.3.- <u>Recomendaciones del uso del cribado en los trastornos depresivos.</u>	63
2.3.1.-Recomendaciones de la <i>United States Task Force</i> (USPSTF).	63
2.3.2.-Recomendaciones de la <i>Canadian Task Force on Preventive Health Care</i> (CTFPHC).	65
2.3.3.- Recomendaciones de la guía del <i>National Institute of Health and Clinical Excellence</i> (NICE).	67
2.3.4.-Recomendaciones de otras organizaciones o instituciones científico-profesionales.	68
2.4.- <u>Controversias en el uso del cribado: resultados meta-analíticos.</u>	70

CAPITULO 3 : INSTRUMENTOS DE CRIBADO DE LOS TRASTORNOS

DEPRESIVOS EN ATENCIÓN PRIMARIA.	79
3.1.- <u>Instrumentos de cribado de los trastornos depresivos en España.</u>	80
3.1.1.-Escala de depresión de Hamilton (HDRS).	83
3.1.2.-Escala de depresión de los estudios epidemiológicos (CES- D.)	91
3.1.3.-Test de 2 preguntas.	93
3.1.4.-Escala autoaplicada para la medida de la depresión de Zung (SDS).	95
3.1.5.-Evaluación de los trastornos mentales en atención primaria (PRIME-MD)	98
3.1.5.1.-Cuestionario de la salud del paciente-9 (PHQ-9).	101
3.1.5.2.-Cuestionario de la salud del paciente-8 (PHQ-8).	105

3.1.6.-Escala Hospitalaria en Ansiedad y Depresión (HAD).	108
3.1.7.-Inventario de la depresión de Beck II (BDI-II).	114
3.1.8.-Escala de Ansiedad y Depresión de Goldberg (EADG).	120
CAPITULO 4: HISTÓRICO EN EL INSTRUMENTO BDI.	125
4.1.- <u>El Inventario de Depresión de Beck (BDI).</u>	126
4.2.- <u>El BDI original (BDI-I) y su adaptación española.</u>	128
4.3.- <u>El BDI-I corregido (BDI-IA) o BDI de 1978 y su adaptación española.</u>	130
4.4.- <u>La segunda edición del BDI (BDI-II) y su adaptación española.</u>	138
4.5.- <u>El Inventario de Depresión de Beck–Cribado Rápido (BDI-FS).</u> <u>o BDI–FS).</u>	140
<u>PARTE EMPÍRICA:</u>	151
CAPÍTULO 5: OBJETIVOS DE LA INVESTIGACIÓN EMPÍRICA.	151
5.1.- <u>Objetivos e hipótesis del primer estudio.</u>	153
5.2.- <u>Objetivos e hipótesis del segundo estudio.</u>	158
CAPITULO 6: <u>PRIMER ESTUDIO:</u> ANÁLISIS DE LAS PROPIEDADES PSICOMÉTRICAS DEL BDI-FS EN ATENCIÓN PRIMARIA EN ESPAÑA.	163
6.1.- <u>Introducción.</u>	164
6.2.- <u>Método.</u>	165
6.2.1.-Participantes.	165
6.2.2.-Instrumentos.	171
6.2.3.-Procedimiento.	175
6.2.4.-Diseño y Análisis de datos.	177
6.3.- <u>Resultados.</u>	180
6.3.1.-Prevalencia de los trastornos depresivos.	180

6.3.2.-Validez Factorial.	181
6.3.3.-Distribución de las puntuaciones del BDI FS.	183
6.3.4.-Fiabilidad de consistencia interna.	186
6.3.5.-Validez Convergente.	187
6.3.6.-Validez Discriminante.	188
6.3.7.-Validez de Criterio.	191
6.4.- <u>Discusión</u>	203
6.4.1.-Validez Factorial del BDI FS.	203
6.4.2.-Sensibilidad del BDI FS.	204
6.4.3.-Fiabilidad del BDI FS.	206
6.4.4.-Validez Convergente.	210
6.4.5.-Validez Discriminante.	211
6.4.6.-Validez de Criterio del BDI FS: rendimiento diagnóstico.	214
6.4.7.-Limitaciones del estudio 1.	218

CAPITULO 7: SEGUNDO ESTUDIO: REPLICACIÓN Y AMPLIACIÓN

DE LAS PROPIEDADES PSICOMÉTRICAS DEL BDI-FS EN	
ATENCIÓN PRIMARIA EN ESPAÑA.	221
7.1.- <u>Introducción.</u>	222
7.2.- <u>Método.</u>	226
7.2.1.-Participantes.	226
7.2.2.-Instrumentos.	230
7.2.3.-Procedimiento.	236
7.2.4.-Diseño y análisis de datos.	238
7.3.- <u>Resultados.</u>	240
7.3.1.-Distribución de las puntuaciones del BDI FS.	240
7.3.2.-Fiabilidad	243
7.3.2.1.-Fiabilidad de consistencia interna del BDI FS	243

7.3.2.2.-Fiabilidad test –retest.	244
7.3.3.-Validez Convergente.	245
7.3.4.-Validez Discriminante.	245
7.3.5.-Validez Factorial.	257
7.3.6.-Diferencias del BDI FS en función de las características demográficas.	260
7.3.7.-Baremos.	262
7.4.- <u>Discusión.</u>	264
7.4.1.-Sensibilidad del BDI FS.	264
7.4.2.-Fiabilidad del BDI FS.	266
7.4.3.-Validez Convergente.	269
7.4.4.-Validez Discriminante.	270
7.4.5.-Validez Factorial.	273
7.4.6.-Diferencias del BDI FS en función de las características demográficas.	274
7.4.7.-Baremos.	276
7.4.8.-Limitaciones del estudio 2º.	278
CAPITULO 8: CONCLUSIONES GENERALES	281
<u>REFERENCIAS BIBLIOGRÁFICAS</u>	295
Referencias bibliográficas.	296
<u>ANEXOS</u>	323
1.1.- <u>Anexos del estudio 1º</u>	324
1.1-Consentimiento informado.	324
1.2.-Protocolo de evaluación.	325
1.2.1.-Datos sociodemográficos.	325

1.2.2.-Documento diagnóstico del médico de atención primaria.	326
1.2.3.-Inventario de Depresión de Beck-Cribado Rápido (BDI FS).	327
1.2.4.-Escala Hospitalaria de Ansiedad y Depresión (HAD).	328
 2.- <u>Anexos del estudio 2º</u>	331
2.1.-Consentimiento informado.	331
2.2.-Protocolo de evaluación.	332
2.2.1.-Datos sociodemográficos.	332
2.2.2.-Inventario de Depresión de Beck-Cribado Rápido (BDI FS).	333
2.2.3.-Escala Autoaplicada para la Depresión de Zung (SDS).	334
2.2.4.-Inventario de Ansiedad de Beck (BAI).	335

ÍNDICE DE TABLAS:

CAPITULO 1: TRASTORNOS DEPRESIVOS EN ATENCIÓN PRIMARIA.

Tabla 1.1 Síntomas que incluyen los criterios diagnósticos de los trastornos depresivos del DSM-IV y de la CIE-10	8
Tabla 1.2 Síntomas asociados a las especificaciones de los trastornos depresivos del DSM-IV y de la CIE-10	11
Tabla 1.3.-Resumen de estudios epidemiológicos en población general.	19
Tabla 1.4.-Prevalencia general en América Latina.	22
Tabla 1.5.-Datos de la prevalencia actual o al año del episodio depresivo mayor en muestras españolas de atención primaria.	28
Tabla 1.6.-Sintomatología depresiva como base del diagnóstico en atención primaria.	47

CAPÍTULO 2: USO DE INSTRUMENTOS DE CRIBADO O DESPISTAJE (SCREENING) DE LOS TRASTORNOS DEPRESIVOS.

Tabla 2.1.-Resumen de las recomendaciones de la Canadian Task Force on Preventive Health Care (CTFPHC) respecto al <i>screening</i> en depresión.	66
Tabla 2.2.-Resumen de los principales resultados sobre cribado.	75

CAPITULO 3 : INSTRUMENTOS DE CRIBADO DE LOS TRASTORNOS DEPRESIVOS EN ATENCIÓN PRIMARIA.

Tabla 3.1.-Puntos de corte del instrumento de depresión Hamilton.	84
Tabla 3.2.-Puntos de corte del instrumento de depresión de Hamilton Fukurawa (2010).	85
Tabla 3.3.-Índices de fiabilidad HDRS Bobes et al. (2003).	86

Tabla 3.4.-Niveles de depresión del HDRS recomendadas por diferentes autores.	90
Tabla 3.5.-Puntos de corte SDS Conde y Chamorro (1973).	96
Tabla 3.6.-Puntos de corte SDS Aragones, Masdeu, Guasch y Borrás (2001).	97
Tabla 3.7.-Puntos de corte PHQ_9 Kroenke et al. (2001).	102
Tabla 3.8.-Puntos de corte PHQ_8 Kroenke et al. (2009).	107
Tabla 3.9.-Puntos de corte HAD Caro e Ibáñez (1992).	110
Tabla 3.10.-Comparación validez convergente y discriminante HAD con otros instrumentos de cribado (Bjelland et al., 2002).	113
Tabla 3.11.-Puntos de corte BDI II.	116
Tabla 3.12.-Puntos de corte para la EADG (Goldberg, 1989).	121
Tabla 3.13.-Cuadro resumen de los principales cuestionarios de cribado en depresión.	123

CAPÍTULO 4: HISTÓRICO DEL INSTRUMENTO BDI.

Tabla 4.1.-Puntos de corte del BDI .Beck , Epstein, Brown, y Steer, (1988).	129
Tabla 4.2.-Puntos de corte del BDI IA Beck y Steer (1993).	131
Tabla 4.3.-Puntos de corte del BDI-IA-SCA Sanz y García Vera (2007).	134
Tabla 4.4.-Puntos de corte del BDI-IA-SF Sanz y García Vera (2007).	136
Tabla 4.5.-Puntos de corte del BDI-FS Beck et al. (2000).	142
Tabla 4.6.-Resumen de las propiedades psicométricas del BDI FS en la literatura científica.	144
Tabla 4.7.-Comparativa BDI IA; BDI II y BDI FS	147
Tabla 4.8.-Resumen validez de contenido BDI IA; BDI II y BDI FS	149

CAPITULO 6: PRIMER ESTUDIO: ANÁLISIS DE LAS PROPIEDADES PSICOMÉTRICAS DEL BDI-FS EN ATENCIÓN PRIMARIA EN ESPAÑA.

Tabla 6.1.-Características sociodemográficas de la muestra de pacientes de atención primaria del Estudio 1º.	167
Tabla 6.2.-Características clínicas de la muestra de pacientes de atención primaria del Estudio 1º.	170
Tabla 6.3.-Autovalores y porcentajes de varianza de los factores extraídos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 1º.	181
Tabla 6.4.-Matriz factorial del BDI-SF en la muestra de pacientes del Estudio 1º tras la extracción de un factor mediante el método de ejes principales.	182
Tabla 6.5.-Distribución de la muestra de pacientes de atención primaria del Estudio 1º en función de los niveles de gravedad del BDI-FS.	185
Tabla 6.6.-Características psicométricas de los ítems del BDI-FS en la muestra de pacientes de atención primaria del Estudio 1º (N = 295).	186
Tabla 6.7.-Correlaciones entre las medidas de depresión y ansiedad en la muestra de pacientes de atención primaria del Estudio 1º.	188
Tabla 6.8.-Análisis factorial conjunto de los ítems del BDI-FS y de la HAD: solución bifactorial.	191
Tabla 6.9.-Índices de rendimiento diagnóstico para los puntos de corte del BDI-FS que obtuvieron los mejores índices en la muestra de pacientes de atención primaria del Estudio 1º.	195
Tabla 6.10.-Índices de rendimiento diagnóstico para el juicio clínico del médico de atención primaria (AP) y para el mejor punto de corte del BDI-FS y de la subescala de depresión de la HAD (HAD-D).	202

CAPITULO 7: SEGUNDO ESTUDIO: REPLICACIÓN Y AMPLIACIÓN DE LAS PROPIEDADES PSICOMÉTRICAS DEL BDI-FS EN ATENCIÓN PRIMARIA EN ESPAÑA.

Tabla 7.1.-Características demográficas de la muestra de pacientes de atención primaria del Estudio 2°.	228
Tabla 7.2. Distribución de la muestra de pacientes de atención primaria del Estudio 2° en función de los niveles de gravedad del BDI-FS	242
Tabla 7.3.-Características psicométricas de los ítems del BDI-FS en la muestra de pacientes de atención primaria del Estudio 2° ($N = 198$).	248
Tabla 7.4.-Análisis factorial conjunto de los ítems del BDI-FS, SDS y BAI: matriz de configuración de la solución bifactorial.	251
Tabla 7.5.-Análisis factorial conjunto de los ítems del BDI-FS, SDS y BAI: matriz de configuración de la solución trifactorial.	256
Tabla 7.6.-Análisis factorial conjunto de los ítems del BDI-FS y BAI: matriz de configuración de la solución bifactorial.	258
Tabla 7.7.-Autovalores y porcentajes de varianza de los factores extraídos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 2°.	258
Tabla 7.8.-Matriz factorial del BDI-SF en la muestra de pacientes del Estudio 2° tras la extracción de un factor mediante el método de ejes principales.	261
Tabla 7.9.-Puntuaciones medias, desviaciones típicas (<i>DT</i>) y puntuaciones medias ajustadas en el BDI-FS de la muestra conjunta de los Estudios 1° y 2° en función de sus características demográficas y resultados de los ANCOVA unifactoriales sobre las puntuaciones en el BDI-FS.	263
Tabla 7.10.-Baremos del BDI-FS para pacientes españoles de atención primaria obtenidos a partir de la muestra conjunta de los Estudios 1° y 2°.	263

ÍNDICE DE FIGURAS:

CAPITULO 1: TRASTORNOS DEPRESIVOS EN ATENCIÓN PRIMARIA:

Figura 1.1. Clasificación de los trastornos depresivos.	17
---------------------------------------------------------	----

CAPÍTULO 2: USO DE INSTRUMENTOS DE CRIBADO O DESPISTAJE (SCREENING) DE LOS TRASTORNOS DEPRESIVOS.

Figura 2.1 Diferentes niveles de prevención de la enfermedad.	58
---------------------------------------------------------------	----

CAPITULO 6: PRIMER ESTUDIO: ANÁLISIS DE LAS PROPIEDADES PSICOMÉTRICAS DEL BDI-FS EN ATENCIÓN PRIMARIA EN ESPAÑA.

Figura 6.1. Resumen del procedimiento del Estudio 1º.	177
-------------------------------------------------------	-----

Figura 6.2. Gráfico de sedimentación de Cattell de los factores obtenidos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 1º.	183
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Figura 6.3. Distribución de las puntuaciones del BDI-FS en la muestra de pacientes de atención primaria del Estudio 1º.	184
----------------------------------------------------------------------------------------------------------------------------	-----

Figura 6.4. Curva ROC o de rendimiento predictivo de las puntuaciones del BDI-FS en la muestra de pacientes de atención primaria del Estudio 1º.	194
--------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Figura 6.5. Curva ROC o de rendimiento predictivo de las puntuaciones de la subescala de depresión de la HAD en la muestra de pacientes de atención primaria del Estudio 1º.	198
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Figura 6.6. Curva ROC o de rendimiento predictivo del diagnóstico del médico de atención primaria en la muestra de pacientes de atención primaria del Estudio 1º.	199
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Figura 6.7. Comparación de las curvas ROC del BDI-FS, de la subescala de depresión de la HAD y del diagnóstico del médico de atención primaria	200
------------------------------------------------------------------------------------------------------------------------------------------------	-----

CAPITULO 7: SEGUNDO ESTUDIO: REPLICACIÓN Y AMPLIACIÓN DE LAS PROPIEDADES PSICOMÉTRICAS DEL BDI-FS EN ATENCIÓN PRIMARIA EN ESPAÑA.

Figura 7.1. Resumen del procedimiento del segundo estudio.	237
Figura 7.2. Distribución de las puntuaciones total del BDI-FS en la muestra de pacientes de atención primaria del Estudio 2º.	241
Figura 7.3. Gráfico de sedimentación de Cattell de los factores obtenidos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS, la SDS y el BAI de la muestra de pacientes de atención primaria del Estudio 2º.	250
Figura 7.4. Gráfico de sedimentación de Cattell de los factores obtenidos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 2º.	259

AGRADECIMIENTOS:

La presente tesis supone la culminación de un trabajo intenso que inicié en el curso 2004/2005 cuando comencé el programa de doctorado denominado “Psicología Clínica Experimental” perteneciente al Departamento de Personalidad, Evaluación y Tratamientos Psicológicos I de la Facultad de Psicología en la Universidad Complutense de Madrid. El trabajo se enmarca dentro de la línea de investigación en la depresión, en este caso en concreto su evaluación, en la cual, el Profesor Sanz lleva veinte años trabajando. Aprovecho para expresarles mi más sincero agradecimiento al Profesor Jesús Sanz y a la Profesora María de la Paz García Vera por acogerme, guiarme, enseñarme y esencialmente hacerme partícipe de su buen hacer en el trabajo. En lo personal gracias por alimentar mis inquietudes y pese a la distancia seguir contando conmigo.

Mi agradecimiento a Pearson Educación, S. A., a través de su división Pearson Clinical and Talent Assessment España, por el permiso para utilizar el BDI-FS y el BAI en la investigación empírica de esta tesis doctoral.

También quisiera mostrar mi gratitud a las personas que han ayudado a que mi proyecto se materializara, facilitándome el trabajo a realizar y coordinándose cuando era necesario, al Dr. Sopelana, al Dr. Fernández Liria y al Dr. Ruiz, gracias por creer en mí desde el principio. Muy especialmente agradecer al grupo de médicos de atención primaria del Centro de Salud “Reyes Magos” de Alcalá de Henares así como a los médicos de primaria de los centros de Salud “Jardinillos”, “La Puebla”, “Eras del Bosque” y “Pintor Oliva” quienes amablemente participaron y apoyaron la investigación. De nuevo, muchas gracias por vuestro tiempo y dedicación. A la gerencia y

comisión de docencia e investigación del “Hospital Universitario Príncipe de Asturias” y del “Complejo Asistencial de Palencia” por permitir que la investigación se llevara a cabo. A Blanca por su efectividad en la recepción de referencias bibliográficas.

También quería hacer mención a mis compañeros de Salud Mental, tanto de Alcalá de Henares como de Palencia, por sus gran apoyo a lo largo de todo el proceso. Especialmente a Violeta por las largas tardes juntas durante la fase del muestreo y a Ruth por darme ánimos cuando se hacía difícil ver el final.

A mis compañeras de departamento, Inés, Regina y María por su necesaria ayuda desde la distancia y porque pese a llevar diferentes caminos, espero sigamos compartiendo inquietudes profesionales y personales.

Este trabajo no hubiera sido posible sin ti, Sergio, gracias de corazón por tu aliento, paciencia y por conseguir que mis ilusiones se cumplan.

Expresar mi reconocimiento a Fran, Mariví, Salva y Puri, por su apoyo incondicional y por las tardes de estudio compartidas. A Emilio y Mayte por sus “visitas fugaces” y siempre estar cuando los necesito y en general a mi familia y amigos, especialmente a Sofía, Vane, Laura, Vero y Rubén por tener la fortuna de tenerlos siempre a mi lado.

Finalmente y fundamentales, a los pacientes y participantes de la investigación, quienes de manera anónima y desinteresada dedicaron su tiempo a mi proyecto.

PRÓLOGO

La evaluación es una parte fundamental de la psicología clínica y de la salud y, en general, de todas las disciplinas clínicas. La evaluación se desarrolla y es importante a lo largo de todo el proceso terapéutico, y así es entendida en esta tesis doctoral, es decir, se trataría tanto de la detección, exploración y diagnóstico de patologías como de la monitorización del progreso terapéutico y de los cambios clínicamente significativos, y para cada uno de esos objetivos es importante contar con instrumentos con las debidas garantías de fiabilidad y validez. En particular, esta tesis doctoral versa sobre la validación psicométrica de un instrumento en la población española de pacientes de atención primaria (AP). Se trata de un instrumento para la evaluación de la sintomatología depresiva y el cribado de trastornos depresivos en pacientes médicos, el Inventario de Depresión de Beck–Cribado Rápido o BDI-FS (*Beck Depression Inventory–Fast Screen*). Sin embargo, este instrumento puede también utilizarse en otros tipos de pacientes (p. ej., pacientes médicos hospitalizados, pacientes geriátricos) y para otros tipos de objetivos (p. ej., para la monitorización del progreso terapéutico) y, sin duda, los resultados de esta tesis doctoral proporcionarán una base sólida para investigar la utilización del BDI-FS en otros tipos de pacientes españoles y con otros objetivos.

Esta tesis doctoral está dividida en dos partes, una teórica y otra empírica. La primera parte o parte teórica tiene un carácter introductorio y en ella se ha tratado el marco teórico del que surge y se fundamenta la presente

tesis. A su vez, esta parte está dividida en cuatro capítulos. El **primer capítulo** trata sobre la conceptualización y definición de la depresión a lo largo de los años y especialmente en la actualidad. También versa sobre la relevancia de la depresión en términos de su epidemiología, tanto a nivel de la población general como de la población de pacientes de AP. En este sentido, se analiza la frecuencia de la depresión, sus factores de riesgo, sus perfiles sociodemográficos, su curso y su comorbilidad con otras patologías. Posteriormente se analizan la importancia a nivel social, emocional, económico y sanitario de la depresión tanto a nivel de la comunidad como desde el punto de vista de la salud pública. Como aspecto fundamental y que conforma uno de los problemas de los que surge la presente tesis, se analizan las dificultades en el diagnóstico de la depresión por parte de los médicos de AP. Finalmente, se revisan brevemente los diferentes métodos de detección de la depresión que se emplean en AP.

En el **segundo capítulo** se delimita la definición y uso del cribado o despistaje de la depresión en AP. De manera paralela, se analizan las recomendaciones al respecto procedentes de los principales organismos e instituciones sanitarias internacionales. El cribado de la depresión en AP es en la actualidad un tema candente, polémico. Por ello, para finalizar se contempla la situación actual en términos de sus discrepancias y se proporcionan las diferentes perspectivas de las instituciones en diversos contenidos relevantes al cribado.

En el **tercer capítulo** se sistematiza la información sobre los principales instrumentos de cribado de la depresión utilizados en AP.

Para finalizar la parte teórica, con el **cuarto capítulo** se contextualiza el instrumento objeto de validación de esta tesis. De este modo, se trata de conocer el marco conceptual del que parte, su génesis y desarrollo a partir de sus diferentes versiones, así como aspectos funcionales en su aplicación en AP.

A partir del análisis de la literatura científica que conlleva esta primera parte de la tesis, se descubre un importante vacío teórico y práctico en España sobre el problema de la detección de la depresión en AP y, en particular, se descubre la inexistencia de pruebas creadas y validadas empíricamente para el despistaje de dicha patología en España. Para solventar en parte estas deficiencias, se plantea la parte empírica de esta tesis doctoral. De esta forma, el objetivo general de la parte empírica de esta tesis doctoral es examinar el comportamiento psicométrico del BDI-FS en la población de pacientes de AP de España, dado que se trata de un instrumento creado específicamente para la detección de la depresión en AP, que a nivel internacional ha dado resultados muy satisfactorios y que en España se desconocen sus propiedades psicométricas.

En este sentido, la fase empírica de esta tesis doctoral implica la realización de dos estudios instrumentales y comprende cuatro capítulos, del quinto capítulo al octavo.

El **quinto capítulo** detalla los objetivos e hipótesis generales y específicos de la investigación empírica.

En el **sexto capítulo** se presenta el primer estudio sobre las principales propiedades psicométricas del BDI-FS (validez factorial, fiabilidad de consistencia interna, validez convergente, validez discriminante, y validez de

criterio o rendimiento diagnóstico) y que se realizó con una muestra de pacientes de AP de Madrid.

En el **séptimo capítulo** se describe el segundo estudio sobre las propiedades psicométricas del BDI-FS y que se llevó a cabo en una muestra de pacientes de AP de Castilla y León. El objetivo general de este segundo estudio era replicar las propiedades psicométricas analizadas en el estudio primero, así como ampliarlas con el estudio de índices psicométricos anteriormente no considerados.

Cabe reseñar que previamente al inicio de los dos estudios, se solicitó y se recibió la aprobación para la realización de dichos estudios, de las Comisiones de Docencia e Investigación del Hospital Universitario “Príncipe de Asturias” (Alcalá de Henares) en el primer estudio, y del “Complejo Asistencial de Palencia” en el segundo.

Finalmente, en el **octavo capítulo** se presentan las conclusiones generales derivadas del análisis y discusión de los resultados de los dos estudios y, en general, de la presente tesis doctoral. Así mismo, se plantean limitaciones de la presente tesis doctoral y posibles líneas futuras de investigación.

La tesis doctoral termina con una sección en la que se listan las referencias bibliográficas consultadas y con una sección de anexos en la que se reproducen los instrumentos y materiales utilizados en los dos estudios empíricos, con el objetivo de que ambas secciones no sólo sirva para justificar y entender la presente tesis doctoral, sino también para replicar sus hallazgos o profundizar en la temática abordada.

PARTE TEÓRICA

*¡Oh nube de mi oscuridad, que me aíslas, sobrevenida de indecible
manera, inflexible e irremediable! ¡Ay, ay de mí de nuevo! ¡Cómo me
penetran, al mismo tiempo, los pinchazos de estos agujones y el
recuerdo de mis males!”*

Sófocles en “Edipo Rey”

Capítulo 1

Trastornos Depresivos en Atención Primaria

“He aquí mi secreto. Es muy simple: no se ve bien sino con el corazón.

Lo esencial es invisible a los ojos”

Antoine de Saint- Exupéry en “El Principito”.

1.1. Conceptualización y clasificación de los trastornos depresivos

1.1.1. Conceptualización de los trastornos depresivos

Definir los trastornos depresivos es una ardua tarea. No en vano ha sido objeto de importantes discusiones a lo largo de los siglos. En la práctica clínica, e incluso a nivel teórico, muy frecuentemente aparece de manera dudosa el concepto “depresivo”. En ocasiones se entiende como un estado de ánimo, otras como síntoma, a veces aparece como síndrome y en otras ocasiones directamente como cuadro clínico. Si se piensa en el lenguaje popular el mismo término también adopta diferentes significados: puede ser un estado (“estoy deprimido”), un rasgo (“es una persona depresiva”), o un síntoma o trastorno (“tiene una depresión”). Respecto a lo que parece que hay consenso es que se trata de un concepto complejo en cuanto a sus acepciones. No en vano, el constructo depresión entendido como trastorno, es decir, como un conjunto de síntomas relacionados que aparecen y desaparecen conjuntamente y para el cual se han especificado ciertos parámetros de duración, gravedad, curso o historia natural y, además, se han descartado algunas posibles causas, ha ido sufriendo transformaciones a lo largo del tiempo y las diferentes épocas, en especial en cuanto a su distinción de la manía.

Durante muchos años se entendió la enfermedad mental en general y los trastornos depresivos en particular como consecuencia de una alienación externa, producto de un agente diabólico, maligno. Parece que los primeros en describir la melancolía desde la propia persona (como un exceso de bilis negra) fueron los autores del *Corpus Hippocraticum* (siglo IV a. C.), atribuidos a los

discípulos de Hipócrates. Éstos, entre los que se encontraba Galeno, consideraban la melancolía como una entidad compuesta por síntomas físicos y mentales originada por una alteración en el equilibrio de dos humores, la flema y la bilis. Posteriormente, Areteo de Capadocia (siglo I a. C.) asocia la melancolía y la manía dentro del mismo continuo: “en otros surgen pura ira y tristeza y un profundo abatimiento, y a éstos los denominados melancólicos [...] me parece que la melancolía y la manía es comienzo y parte de la manía, ya que el sentido del maniaco tiende a veces a la ira y otras al regocijo”. Continuando en Grecia, Platón diferencia entre la manía patológica y la manía divina, siendo esta última la más elevada *simetría* de lo genial en armonía con el cosmos “el momento del movimiento desde la medida equilibrada a lo desmesurado, hipérbole”. El estado psíquico que coincide con ello es designado como frenesí, manía. Aristóteles considera los trastornos afectivos como fríos “la bilis negra, que por naturaleza es fría...puede...cuando excede en el cuerpo la justa medida, provocar parálisis, rigideces, depresiones o estados de ansiedad. Pero si se calienta en exceso, provoca alborozo de modo tal que se canta y éxtasis”.

Durante el Renacimiento la sociedad entiende a los locos como poseídos y se les persigue cruelmente. Paulatinamente se va superando la teoría de los humores. Autores como R. Burton entienden que existen factores etiológicos diferentes y en 1621 publica “*The anatomy of the melancholy*” donde aparecen explicaciones de carácter psicológico en los trastornos afectivos. Progresivamente se fueron entendiendo de una manera más humanística y científica los trastornos mentales. Dreyssig (1770-1880) formula la primera clasificación racional de los enfermos mentales y distingue tres

formas básicas: la melancolía o locura parcial (subdividida a su vez en verdadera y falsa: apatía e indiferencia), la manía y la imbecilidad. En Francia, Esquirol (1772-1840) incluye la melancolía en la lipemanía (estados depresivos en general), diferenciándola de la manía, la demencia, la idiocia, la imbecilidad y la monomanía (paranoia).

Falret en 1851 describe la melancolía periódica y la locura circular, caracterizada por tres estadios bien diferenciados: una fase melancólica, una fase maniaca y un intervalo lúcido de duración variable. La escuela alemana, y en particular Kraepelin (1856-1926), cuya herencia en la concepción actual de la salud mental ha sido fundamental, clasifica y define, sobre la base de los síntomas, el pronóstico y los procesos subyacentes, los diferentes tipos de enfermedades mentales, y distingue tres estados afectivos: la melancolía, la locura circular o periódica y la manía simple. En 1896, Kraepelin introdujo el concepto de locura maniaco-depresiva y en 1899 describe la melancolía involutiva. Heinroth en 1918 utiliza la palabra depresión como sinónimo de astenia. La corriente psicoanalítica, liderada por Freud (1900/1953), preponderante en las décadas de los 20-50 del siglo pasado, se centra en la influencia de las ideas inconscientes sobre la emociones. K. Snneider en 1963 acuña el término tristeza vital para referirse a la tristeza del depresivo de diferente cualidad. (Reseñas históricas tomadas de Chinchilla, 2008; Eguiluz, 2001).

Desde finales del siglo pasado y, en particular desde la publicación de la tercera edición del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-III) de la Asociación Americana de Psiquiatría (APA, 1980), el concepto de depresión viene marcado por la hegemonía de las clasificaciones

nosológicas, las cuales han tratado de consensuar, operativizar y dar un orden a las observaciones, conceptos y cuadros clínicos que se habían ido creando a lo largo de la historia. Creadas desde una perspectiva ateórica y con un marcado sentido categorial, las clasificaciones nosológicas a partir de los años 70-80 del siglo pasado han facilitado la comprensión de la salud mental en general y, de los trastornos depresivos en particular, al crear un lenguaje común para todas las orientaciones teóricas sobre la psicopatología.

De manera resumida se puede observar como el concepto de trastorno depresivo ha ido variando a lo largo de los años. Actualmente, los conceptos de trastorno depresivo más consensuados y validados, aunque no lejos de polémica y discusión, son los propuestos por la cuarta edición del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-IV) de la Asociación Americana de Psiquiatría (APA, 1994) y la clasificación de trastornos mentales y del comportamiento de la décima revisión de la Clasificación Internacional de Enfermedades (CIE-10) de la Organización Mundial de la Salud (OMS, 1993), los cuales distinguen a los trastornos depresivos de los trastornos del estado de ánimo que incluyen la manía o hipomanía (trastornos bipolares) y, a su vez, dentro de los trastornos depresivos distinguen al trastorno depresivo mayor y al trastorno distímico. Además, los conceptos de trastorno depresivo del DSM-IV y la CIE-10 guardan muchas semejanzas entre sí respecto a los síntomas y parámetros de duración, gravedad, curso e historia natural que definen el trastorno depresivo mayor y la distimia (véase la Tabla 1.1).

Tabla 1.1. Síntomas que incluyen los criterios diagnósticos de los trastornos depresivos del DSM-IV y de la CIE-10 (tomado de Sanz, en prensa)

Síntomas	DSM-IV (1994)		CIE-10 (1993)	
	Depresión Mayor	Distimia	Depresión Mayor	Distimia
Anímicos				
1. Estado de ánimo o humor deprimido *	X	X	X	X
13. Llanto fácil				X
Motivacionales				
2. Disminución-pérdida del placer o interés	X		X	X
Físicos				
3. Fatigabilidad o pérdida de energía o vitalidad	X	X	X	X
4. Aumento/disminución del peso o apetito	X	X	X	
5. Insomnio o hipersomnia o alteraciones del sueño	X	X	X	X
Motores				
6. Agitación o enlentecimiento psicomotor	X		X	
9a. Conducta o intentos suicidas †	X		X	
15. Aislamiento social				X
16. Disminución de la locuacidad				X

Nota. † El criterio sintomático 9 de actos/ideación suicidas se ha considerado como dos síntomas, uno motor (9a: actos suicidas) y otro cognitivo (9b: ideación suicida).

* En el DSM-IV también se incluye la posibilidad de que en niños y adolescentes el estado de ánimo pueda ser irritable.

Tabla 1.1. Síntomas que incluyen los criterios diagnósticos de los trastornos depresivos del DSM-IV y de la CIE-10 (tomado de Sanz, en prensa) (Continuación)

Síntomas	DSM-IV (1994)		CIE-10 (1993)	
	Depresión Mayor	Distimia	Depresión Mayor	Distimia
Cognitivos				
7. Sentimientos excesivos o inapropiados de inutilidad o culpa o auto-reproches desproporcionados	X		X	
8. Disminución de la capacidad para pensar, concentrarse o tomar decisiones	X	X	X	X
9b. Ideas recurrentes de muerte o suicidio †	X		X	
10. Pérdida de confianza y autoestima, y sentimientos de inferioridad		X	X	X
11. Sentimientos de desesperanza o desesperación		X		X
12. Sentimiento de incapacidad para enfrentarse con las responsabilidades de la vida diaria				X
14. Pesimismo sobre el futuro o cavilaciones sobre el pasado				X

Nota. † El criterio sintomático 9 de actos/ideación suicidas se ha considerado como dos síntomas, uno motor (9a: actos suicidas) y otro cognitivo (9b: ideación suicida).

* En el DSM-IV también se incluye la posibilidad de que en niños y adolescentes el estado de ánimo pueda ser irritable.

Así, por ejemplo, ambos conceptos comparten la idea de que los dos síntomas principales de los trastornos depresivos son el bajo estado de ánimo y la anhedonia, que la duración mínima de los síntomas depresivos debe ser al menos dos semanas en el caso del trastorno depresivo mayor o dos años en el caso del trastorno distímico, y que los síntomas deben provocar en la persona malestar clínicamente significativo o dificultades para llevar a cabo su actividad laboral y social o deterioro de otras áreas importantes de la actividad del individuo (CIE-10, OMS, 1993). Es más, tal y como se puede observar en las Tablas 1.1 y 1.2, los dos conceptos muestran muchas coincidencias en los síntomas definitorios más allá del estado de ánimo deprimido y de la anhedonia (p. ej., fatigabilidad o pérdida de energía o vitalidad, aumento o disminución del peso o apetito, alteraciones del sueño, ideación y conductas suicidas) y en los síntomas asociados que conforman distintas especificaciones del trastorno.

Teniendo en cuenta ambas concepciones del trastorno depresivo, la del DSM-IV y la de la CIE-10, se puede afirmar que, respecto a los síntomas que definen tales trastornos a nivel afectivo, la persona con un trastorno depresivo siente tristeza, infelicidad, pesadumbre, vacío, abatimiento durante la mayor parte del día. Presenta un reducido o nulo disfrute por aspectos con los que antes sentía placer y un bloqueo en la capacidad de autorrecompensarse con estímulos que anteriormente sí lo hacían. También siente apatía, baja motivación por iniciar nuevas conductas, e indiferencia afectiva que incluso llega a convertirse en anestesia a nivel emocional: un paradójico “sentir un no sentir” que genera un gran malestar. También aparece habitualmente la tendencia al llanto sin desencadenante aparente o una hiperreactividad al estímulo precipitante (Eguiluz, 2001).

Tabla 1.2. Síntomas asociados a las especificaciones de los trastornos depresivos del DSM-IV y de la CIE-10 (tomado de Sanz, en prensa)

DSM-IV (1994)	CIE-10 (1993)
Síntomas psicóticos:	Síntomas psicóticos:
- Ideas delirantes	- Ideas delirantes
- Alucinaciones	- Alucinaciones
Síntomas catatónicos:	
- Inmovilidad motora	- Estupor depresivo (disminución o ausencia total de movimientos espontáneos con mutismo parcial o total, negativismo y posturas rígidas mantenidas)
- Actividad motora excesiva	
- Negativismo extremo	- Estupor depresivo
- Peculiaridades del movimiento voluntario	
- Ecolalia o ecopraxia	
Síntomas melancólicos:	Síntomas somáticos:
- Falta de reactividad a los estímulos habitualmente placenteros	- Pérdida de reactividad a los estímulos que habitualmente producen una respuesta emocional
- Calidad distintiva del estado de ánimo depresivo	
- Empeoramiento por la mañana	- Empeoramiento matutino
- Despertar precoz	- Despertar 2 horas antes de los usual o más
	- Pérdida marcada de la libido
Síntomas atípicos:	
- Reactividad del estado de ánimo	
- Abatimiento (sentir brazos o piernas pesados o inertes)	
- Sensibilidad al rechazo interpersonal	

A nivel social, se produce un aislamiento del entorno. Se reduce la comunicación a nivel interpersonal y hay un lenguaje bradipsíquico, con discurso tendente a la parquedad y respuestas lacónicas.

A nivel cognitivo en los casos más graves aparecen alteraciones de la conciencia, llegándose a mostrar el paciente depresivo confuso o incluso perplejo. También aparece una dificultad en la capacidad asociativa y una tendencia a la anticipación y perseveración en contenidos negativos del pensamiento. La atención se dirige a uno mismo y la concentración se ve disminuida y son frecuentes las quejas acerca de la memoria, confundiéndose con las denominadas pseudodemencias depresivas. La percepción del depresivo torna a gris, sin vida. Son habituales en la depresión las anticipaciones negativas, distorsiones cognitivas, errores de pensamiento, autorreproches y culpabilización. En las depresiones con síntomas psicóticos los contenidos delirantes suelen versar sobre ruina, nihilismo, incapacidad, catastrofizaciones..., y habitualmente son congruentes con el estado afectivo del depresivo. En los casos más graves aparecen deseos de suicidio que pueden llegar a convertirse en actos y conllevar el fallecimiento del paciente.

A nivel somático aparecen sentimientos de anergia y fatiga, una disminución de la libido sexual e insomnio de inicio o despertar precoz. También puede aparecer hipersomnia con tendencia a la clinofilia y aumento o disminución del apetito con su consecuente variación ponderal. A nivel psicomotriz aparece frecuentemente una inhibición o un enlentecimiento en la realización de los comportamientos habituales, y, con menos frecuencia, agitación psicomotriz. La persona presenta una postura habitual de repliegue, con facies hipomímica. Suele aparecer de forma asociada un estado de

ansiedad con importante cortejo vegetativo y múltiples quejas somáticas (gastrointestinales, cefaleas tensionales, dolor precordial, opresión torácica, etc.). (Sintomatología tomada del DSM_IV, 1994 y CIE_10, 1992)

1.1.2. Clasificación de los trastornos depresivos

Con la creación de las clasificaciones nosológicas como las del DSMI-IV y la CIE-10, desde el punto de vista clínico se gana en sistematización y claridad al tener un referente en común. Las clasificaciones nosológicas facilitan la comparación entre investigaciones a nivel internacional, lo que a su vez redundará sobre una mejora en el conocimiento de la nosología en salud mental. Al introducir el sistema multiaxial diagnóstico se favorece una comprensión más completa del cuadro clínico, así como supone un esfuerzo por optimizar el entendimiento universal entre profesionales de la salud mental.

Pero con las clasificaciones aparecen también desventajas notables. De hecho, se echa en falta el sentido fenomenológico que subyace a la psicopatología. Específicamente, en los trastornos del humor (anteriormente afectivos, el cual se modifica al eliminar la ansiedad de la entidad), se entienden cuadros clínicos que aunque etiquetados con el mismo nombre, pueden ser muy diferentes respecto a los síntomas concretos que experimentan los pacientes o a la forma en que se experimenta un mismo síntoma. Por ejemplo, frecuentemente se presenta de manera diferente el ánimo depresivo de un trastorno depresivo mayor con etiología reactiva a un suceso externo respecto al ánimo depresivo característico de un anciano, o el de un trastorno bipolar al de una persona cuya depresión tiene una clara base

orgánica. Por otro lado, no parece desdeñable la influencia de la personalidad premórbida en los trastornos afectivos, la cual, se obvia en las actuales clasificaciones diagnósticas. Cabe hacer la misma consideración en relación con la baja especificidad de contenido entre la depresión crónica, la distimia y la personalidad depresiva, la cual hace que los diagnósticos diferenciales entre ellas sean confusos. De ahí que la clasificación de la depresión se complejiza según el criterio de referencia. Aunque no existe un consenso claro sobre la mejor forma de clasificación de los trastornos depresivos, sin pretender ser completamente exhaustivos se muestran los enfoques que clásicamente se han venido utilizando (Fernández, 1999, Sadock y Sadock, 2005):

1. Fundamentado en la etiología:

1.1 Depresión reactiva frente a depresión endógena:

Fundamentalmente relacionado con el factor causal. Se asocian las reactivas a agentes psicológicos o ambientales mientras que en las endógenas habría un desencadenante de predominio más biológico (genético, metabólico).

1.2 Trastorno afectivo primario frente a secundario: El primer término se relaciona con personas que no tienen antecedentes psicopatológicos previos, excepto presencia de episodios de depresiones o manías y en el que no existe comorbilidad de enfermedad física que preceda o evolucione de manera paralela al síndrome afectivo. En el caso contrario se consideraría como secundario.

2. Fundamentado en los síntomas:

2.1 Depresión neurótica frente a psicótica: Esta dicotomía se considera en función de la presencia de síntomas psicóticos como alucinaciones o delirios en el caso de las psicóticas o en las neuróticas de “rasgos de personalidad neuróticos”, entendidos desde lo emotivo, como inestabilidad emocional, una tendencia global a presentar respuestas de ansiedad, a ser vegetativamente hiperreactivo, a mostrar una mayor fatigabilidad física y mental o a tener baja tolerancia a la frustración.

2.2 Depresión atípica: caracterizada por elevada reactividad del estado de ánimo o síntomas no tan habituales en la depresión como hiperfagia, hipersomnia, abatimiento o un patrón de larga duración de sensibilidad al rechazo emocional.

2.3 Depresión melancólica: con síntomas como pérdida de placer en casi todas las actividades, falta de reactividad en estímulos habitualmente placenteros, empeoramiento matutino, despertar precoz, enlentecimiento motor, anorexia, culpabilidad inapropiada o estado de ánimo cualitativamente diferente al habitualmente experimentado.

2.4 Depresión catatónica: Caracterizada por síntomas como estupor, embotamiento afectivo, negativismo, retardo psicomotor con postura y flexibilidad cética.

3. Fundamentada en el curso:

3.1 Patrón estacional: singular por el inicio y remisión del episodio

depresivo mayor en una determinada época del año.

Fundamentalmente los episodios suelen ocurrir en invierno y otoño y remitir en primavera.

3.2 Inicio en el post-parto: Cuando el inicio del episodio depresivo se

presenta en las primeras cuatro semanas tras el alumbramiento de un hijo,

3.3 Relacionada con la evolución longitudinal se distinguen entre

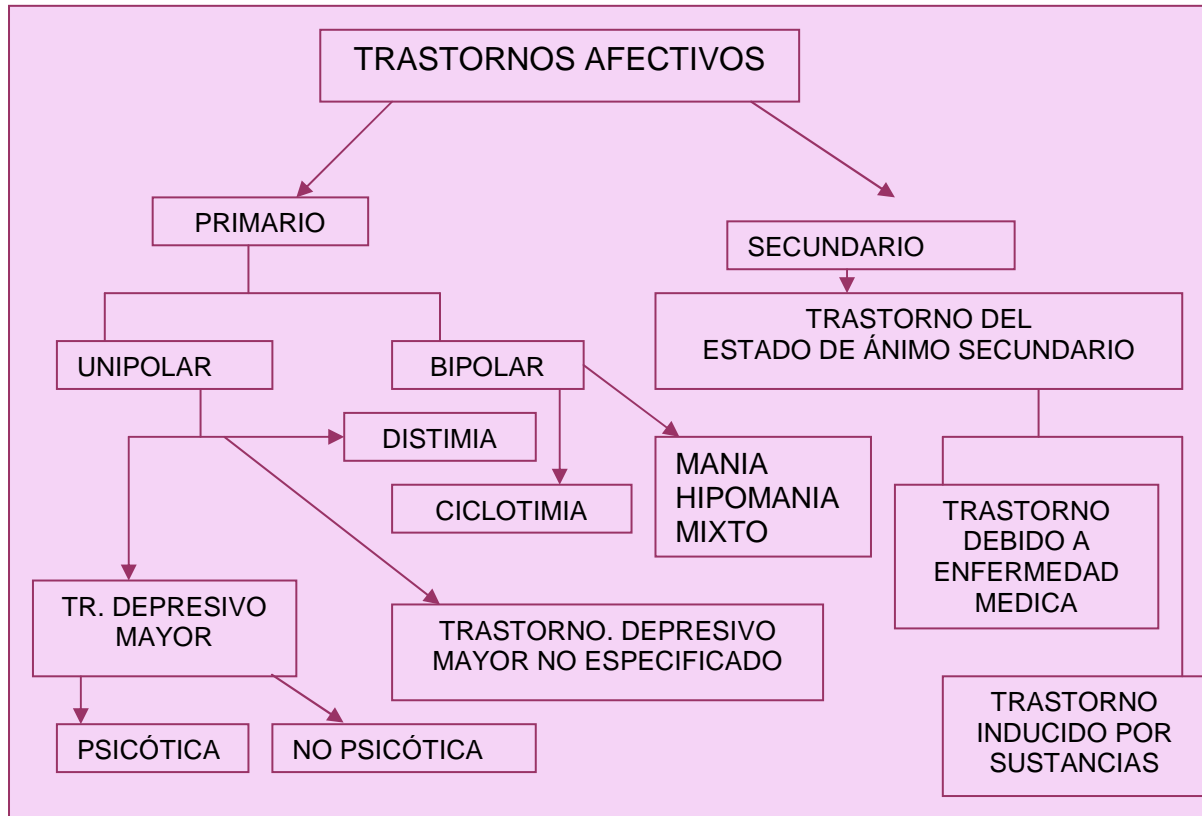
episodios únicos, recidivante, con/sin recuperación interepisódica o en remisión.

Actualmente las clasificaciones de los trastornos afectivos que mayor aval y reconocimiento han obtenidos son, como se ha dicho anteriormente, las procedentes del Manual Diagnóstico y Estadístico de los Trastornos Mentales de la Asociación América de Psiquiatría, en su cuarta edición (DSM-IV; APA, 1994), y el capítulo F (V) sobre trastornos mentales y del comportamiento de la Clasificación Internacional del Enfermedades de la Organización Mundial de la Salud (CIE-10; OMS, 1993).

En términos generales, los trastornos del estado de ánimo están compuestos por aquellos cuya característica principal es la alteración en el humor. Se podría definir el humor como el tono hedónico dominante en un momento determinado, algo así como el color experiencial de un individuo. Los trastornos del estado del ánimo están constituidos por los trastornos

depresivos, los trastornos bipolares y los trastornos basados en su etiología (véase la Figura 1.1).

Figura 1.1. Clasificación de los trastornos afectivos



Nota. Adaptado de Benítez, Casquero, Sánchez, Villalba y Zamorano (2009)

1.2. Prevalencia de la depresión en la población general

Previo a los años 70 del siglo pasado existía una importante dificultad para analizar y generalizar datos epidemiológicos. Aproximadamente a principios de los años setenta del siglo XX se crean los criterios diagnósticos *Feighner* y los de investigación (RDC), los cuales permiten una paulatina operativización de los diferentes trastornos mentales. Paralelamente, se crean las entrevistas estructuradas, permitiendo un marco común a la hora de comparar estudios e investigaciones. Inicialmente se crea la SADS (Entrevista

estructurada para los trastornos afectivos y esquizofrenia), con la ventaja de permitir de una manera sistemática y objetiva explorar este tipo de patología. A nivel epidemiológico tenía dos grandes inconvenientes: la primera, la necesidad de ser realizada por un clínico, y la segunda, que sólo abarcaba determinadas patologías mentales. Posteriormente, con el fin de superar entre otros estos dos inconvenientes, el NIMH (*National Institute of Mental Health*) crea un entrevista diagnóstica basada en criterios del DSM-III (APA, 1980) y llamada DIS (*Diagnostic Interview Schedule*). Haciendo uso del instrumento DIS, apareció a principios de los años 80 del siglo pasado el primer estudio epidemiológico a gran escala de los trastornos mentales, entre ellos la depresión. Se denominó ECA (*Epidemiologic Catchment Area*, 1980-1985) y se realizó en EE UU por parte del NIMH.

Los estudios epidemiológicos proliferaron a lo largo de los años 80 y 90 del siglo pasado, pero muchas veces los datos sobre depresión eran de carácter regional, con muy diferentes metodologías y la posibilidad de generalizaciones eran muy limitadas. Durante los años 90 del siglo XX se hace énfasis en la necesidad de realizar estudios epidemiológicos en los que se puedan comparar datos tanto a nivel nacional como internacional con el fin de hallar conclusiones significativas. Dadas las limitaciones de los instrumentos existentes, la OMS, basándose en los criterios CIE-10, crea la entrevista clínica estructurada CIDI (*Composite International Diagnostic Interview*), la cual, desde el 1990 y en sus diferentes versiones, va a convertirse en un instrumento de referencia común para los diferentes estudios epidemiológicos en depresión. Otras entrevistas ampliamente utilizadas han sido la SCID (Spitzer et al., 1999) basada en criterios DSM-IV y, en menor medida, la SCAN (*Schedule for*

Clinical Assessment in Neuropsychiatry; Wing et al., 1990) o la MINI (*Mini-International Neuropsychiatric Interview*; Sheehan et al., 1998).

Por tanto, si hasta hace unos 15-20 años los datos publicados sobre la prevalencia de los trastornos del estado de ánimo frecuentemente eran confusos o aparecían importantes dificultades a la hora de realizar comparaciones entre las diferentes investigaciones debido a las diferencias en el muestreo, población diana, instrumento de evaluación o análisis de datos, en las últimas dos décadas se han sistematizado y promovido estudios epidemiológicos de mayor alcance, tanto a nivel de tamaño de la muestra como de la inclusión de estudios internacionales, que permiten ir conociendo con mayor precisión la prevalencia de la depresión y las diferencias sociodemográficas y transculturales.

A nivel Europeo el primer estudio epidemiológico sobre la depresión de mayor envergadura fue el denominado DEPRES (*Depression Research in European Society*; Lepine et al., 1997), seguido por su segunda fase DEPRES II (*Depression Research in European Society II*; Tylee et al., 1999); posteriormente, en 2001, se realizó el estudio llamado ODIN (*Outcome of Depression International Network*) coordinado por los profesores españoles Ayuso-Mateos y Vázquez Barquero. Con carácter más comprehensivo aparece el ESEMeD/MHEDEA en el año 2000, que mide no sólo prevalencia sino también factores de riesgo, incapacidad, factores de salud asociados a la calidad de vida, uso de los servicios de salud mental, tratamientos asociados a los trastornos del estado de ánimo, ansiosos y abuso de sustancias en Europa.

Tabla 1.3. Resumen de estudios epidemiológicos sobre la prevalencia de la depresión en población general

<i>Referencia</i>	<i>N</i>	<i>Población</i>	<i>Muestreo</i>	<i>Instrumento</i>	<i>Tipo de prevalencia</i>	% de episodio depresivo mayor		
						<i>Total</i>	<i>En hombres</i>	<i>En mujeres</i>
ESEMED-Spain	5473	España	Transversal	CIDI 3.0	Durante la vida	10,6%	6,4%	14,5%
Gabilondo et al. (2010)			Estratificado.	SCID-I	Últimos 12 meses	4%	2,2%	5,7%
Estudio NCS-R	9090	Estados Unidos de América	Probabilístico	CIDI 3.0	Durante la vida	16,2%	1%	1,7%
Kessler et al. (2003)		(48 estados)	Estratificado	SCID-I	Últimos 12 meses	6,6%	1%	1,4%
			Bifase					
Actualización de ESEMED/MHEDEA 2000	21425	Europa	Transversal	CIDI 3.0	Durante la vida	13,4%	9,4%	17,1%
Alonso et al. (2007)		(Bélgica, Francia, Alemania, Italia, Holanda, España)	Bifase		Últimos 12 meses	4,1%	2,8%	5,3%

Noa. CIDI 3.0 = Entrevista Diagnóstica Internacional Compuesta, versión 3 (Kessler y Ushin, 2004); SCID-I = Entrevista Clínica Estructurada para los Trastornos del Eje I del DSM-IV (First et al., 1995).

En la Tabla 1.3 se resumen las principales características y resultados de los estudios epidemiológicos más importantes y recientes sobre la prevalencia de la depresión en población general.

El estudio americano del NCS-R (*National Comorbidity Survey–Replication*; Kessler et al., 2003), catalogó el 85% de todos los trastornos del estado de ánimo como moderados a graves, interfiriendo sobre la vida de la persona negativamente todos los días (Baumeister y Härter, 2007).

En una revisión realizada en 2005 sobre los trastornos mentales en América Latina y el Caribe, Kohn et al. (2005) encuentran datos sobre epidemiología dispares a lo largo de los diferentes países, lo cual es esperable dadas las diferentes condiciones tamaños muestrales, condiciones sanitarias y condiciones socioeconómicas (véase la Tabla 1.4).

Como ya se ha referido antes, es difícil realizar comparaciones entre países. Intentando equiparar los estudios con mayores equivalencias metodológicas, parece que en EEUU aparecerían las mayores prevalencias sobre depresión. En Europa, Reino Unido sería uno de los países que muestra mayores tasas de depresión. España, Italia y Alemania estarían a la cola. Los datos de América Latina serían más similares a los europeos que a los norteamericanos.

Tabla 1.4. Prevalencia de la depresión en la población general de América Latina

Referencia	Base de datos	Instrumentos	Países	Tipo de prevalencia	% de episodio depresivo mayor		
					Total	En hombres	En mujeres
Kohn et al. (2005) Revisión	MEDLINE	PSE	Brasil	Toda la vida	12,6%	8,8%	15,3%
	LILACS	DIS		12 meses	5,8%	3,6%	7,5%
	Años: 1980-2004	CIDI	Chile	Toda la vida	9,2%	5,7%	11,5%
				12 meses	5,7%	3,7%	7,5%
			Colombia	Toda la vida	12,6%	8,6%	14,9%
				12 meses	5,6%	3,5%	7,3%
			México	Toda la vida	8,1%	5,5%	10,1%
				12 meses	5,6%	3,1%	5,6%
			Perú	Toda la vida	4,6%	3,5%	5,5%
				6 meses	2,3%	2,4%	3,3%
			Puerto Rico	Toda la vida	4,6%	3,5%	5,5%
				6 meses	2,3%	2,4%	3,3%
			Media de todos los países	Toda la vida	8,7%	6%	11,5%
				12 meses	4,9%	3,1%	6,5%

Nota. PSE= Present State Examination; DIS = Diagnostic Interview Schedule; CIDI = Composite International Diagnostic Interview.

Se ha intentado buscar factores etiológicos que expliquen las diferencias entre países en los datos de prevalencia de los trastornos depresivos, como, por ejemplo, la existencia de una cultura más independiente frente a una cultura comunitaria, la influencia de la religión, la cultura mediterránea..., pero los datos no son concluyentes.

A lo largo de los diferentes países, se han encontrado una serie de factores sociodemográficos asociados a depresión de manera consistente: ser mujer, tener un nivel educativo bajo, estar separado, divorciado o viudo, o tener bajo estatus socioeconómico (Alonso et al., 2007; Haro et al., 2005). La franja etaria de mayor probabilidad de padecer un episodio depresivo mayor estaría entre los 30-50 años en hombres y en mujeres de 18-30 años (Ferrer y Rodríguez, 1999). Aparece también una menor tasa de trastorno depresivo en mayores de 65 años (Bellón et al., 2008; King et al., 2008). Otros factores que aparecen de manera sistemática son ser desempleado, dedicarse a las labores de casa y tener incapacidad laboral transitoria, los cuales, aumentarían el riesgo de padecer un cuadro anímico (Alonso et al., 2007; Serrano-Blanco et al., 2008). Un inicio temprano aumentaría la probabilidad de cronicidad del cuadro depresivo (Gabilondo et al., 2010). A nivel asistencial, se encontraría un mayor riesgo en ser hiperfrecuentador de las consultas de AP (Aragonés, Masdéu, Cando y Coll, 2001).

Las enfermedades físicas tendrían un papel en el inicio y en el mantenimiento de los cuadros depresivos. Los acontecimientos vitales estresantes tendrían mayor influencia sobre las mujeres respecto al sexo masculino. El apoyo social y estar casado o conviviendo con una pareja estable

actuarían como factores protectores frente a la depresión. Otros factores menos relacionados serían haber tenido una infancia desfavorecida, consumo de tóxicos y tener ansiedad como enfermedad mental comórbida (Bellón et al., 2008).

Parece que las mujeres tendrían mayor tendencia a presentar cuadros depresivos o de ansiedad y los varones al abuso de sustancias (ESEMeD/MHEDEA 2000; Alonso et al., 2004).

En relación con el curso y los factores predictores de la depresión existen multitud de estudios. Uno de los más relevantes fue el LIDO (*Longitudinal Investigation of Depression Outcome*; Chisholm et al., 2001; Herman et al., 2002). Se trata de un estudio internacional sobre la depresión y sus factores implicados realizado en 6 países (Australia, Brasil, Israel, España, EE. UU. y Rusia) y que forma parte de la iniciativa para el desarrollo y estudio de los instrumentos de calidad de vida de la Organización Mundial de la Salud (WHOQoL). En dicho estudio se señala que la remisión completa de la depresión en los diferentes países a lo largo de los 9 meses era escasa (variando de un 25% en Porto Alegre a un 48% en Barcelona). Como factores asociados al mantenimiento aparecían de manera moderada: nivel educativo, calidad de vida y acontecimientos vitales estresantes (Pío de Almeida et al., 2005).

Otro estudio reseñable es el PREDICT-NL que analiza el curso y evolución de la depresión en AP en Holanda a lo largo de 39 meses. Encontraron que un 57% de las personas deprimidas tenían un curso intermitente o crónico. La baja remisión sintomática estaba asociada con la

gravedad de la depresión, la sintomatología física y la enfermedad mental premórbida. En un estudio realizado en colaboración con la OMS encontraron que tras doce meses, un 66,5% no tenía depresión y un 33,5% seguía cumpliendo criterios según la CIE-10 (Barkow, Maier, Bedirhan, Gäsicke, Wittchen y Heun, 2003). Otros estudios encuentran que un 20% tiene un curso crónico y un 30-50% tiene un curso recurrente (Eaton et al., 2008; Stegenga, Kamphuis, King, Nazareth y Geerlings, 2010).

Cabe reseñar la dificultad para aislar e identificar factores de riesgo en relación con el inicio o mantenimiento de la depresión dada la influencia de multitud de factores previos, concomitantes y premórbidos, lo que complica el control de estos factores a la hora de realizar los estudios epidemiológicos.

1.2.1. Prevalencia de la depresión en atención primaria

La prevalencia de la depresión en AP debe ser diferenciada de la de la población general. En la población general, y en función de los datos presentados en el epígrafe anterior, se podría estimar que la prevalencia es aproximadamente del 6% al año y asciende al 15% a lo largo de la vida. Existen, sin embargo, diferencias en esas cifras en función del sexo: la prevalencia a lo largo de la vida para el episodio depresivo mayor en países industrializados está estimada en el 8-12% en los varones y en el 20-26% en las mujeres. Por otro lado, en todos los países industrializados la incidencia de los trastornos depresivos se está incrementando y estos trastornos están comenzando en edades cada vez más jóvenes (Gabarrón et al., 2002).

En comparación, las cifras de prevalencia de la depresión en AP son ligeramente superiores a las que se encuentran en la población general. Como se puede observar en la Tabla 1.5, la prevalencia al año de la depresión en AP en España se podría estimar en un 13,2%, más del doble de la prevalencia anual en población general que se estimaba en el párrafo anterior (6%).

Se calcula que la depresión supone aproximadamente un 31% de las consultas de AP, siendo mayor el porcentaje de las depresiones que se presentan como síntomas asociados a otros trastornos psíquicos (58%) que como un cuadro depresivo aislado (42%). Además, en AP la depresión se detecta fundamentalmente de forma secundaria (72% de los casos); sólo en el 27% de los casos diagnosticados finalmente de depresión, ésta constituía el motivo inicial de consulta (Martín-Águeda et al., 2005). Los cuadros subclínicos pueden estar presentes en un 10-30% de los pacientes y aproximadamente un 25% de personas con síntomas depresivos no buscan asistencia sanitaria (Gabarrón et al., 2002).

Dada la alta prevalencia de los cuadros depresivos en AP, tal y como se puede observar en la Tabla 1.5, y dada la persistencia de esos cuadros en el tiempo, se hace necesario que para mejorar tanto su detección como tratamiento se conozca el mayor número de factores implicados en su desarrollo y mantenimiento. Coincidente con población general, en los estudios realizados en AP, aparecen datos sociodemográficos asociados a la prevalencia de la depresión de manera consistente. Así, ser mujer, joven, con bajo nivel socioeconómico, estar desempleado, ser soltero o haber estado casado, es decir ser actualmente separado, divorciado o viudo, son factores de riesgo para el inicio de la depresión y su mantenimiento. Otros factores

asociados al mantenimiento de la depresión serían alta gravedad de depresión premórbida, neuroastenia, distimia, agorafobia, dependencia de alcohol, bajas habilidades sociales, presencia de pensamientos de suicidio y no tener estudios o tener menos de 6 años de formación académica (Barcks et al., 2003).

En relación con la prevalencia de la depresión en la población geriátrica hay datos contradictorios. En algunos estudios encuentran que en dicha población hay una menor prevalencia de depresión y que el factor sexo no sería tan relevante y pasarían a ser factores de riesgo vivir en medio rural, tener riesgo social, bajo apoyo familiar, sufrir un accidente cerebrovascular y padecer una discapacidad sensorial (Urbina Torija, 2007). Sin embargo en otros estudios encuentran tasas elevadas de depresión y otros factores de riesgo. El estudio PSICOTARD encuentra un 14.3% de prevalencia e informa que, en su revisión de estudios previos, encuentran variaciones en la prevalencia desde un 0,4% hasta un 35%, asociando dichas diferencias a variaciones metodológicas de muestreo, instrumento, etc. El estudio también encuentra como correlatos de la depresión el sexo femenino, la gravedad de la enfermedad física y vivir solo (Olivera et al., 2008). Barcks et al. (2003) encontraron que en personas mayores de 54 años, en comparación con personas más jóvenes, aparecía una mayor *odds ratio* (OR) en la prevalencia anual de la depresión.

Tabla 1.5. Datos de prevalencia actual o al año del episodio depresivo mayor en muestras españolas de pacientes de atención primaria

Referencia	Nº de pacientes	Muestreo/Diseño	Instrumentos	Prevalencia total	Prevalencia en varones	Prevalencia en mujeres
Gabarrón et al. (2002)	400	Incidental/una fase, pero datos ajustados por rendimiento diagnóstico	BDI MINI	20,2%	8,1%	26,8%
Aragonés et al. (2001)	350	Incidental/doble fase	SDS SCID	14,7%	–	–
Aragonés et al. (2004)	906	Incidental/doble fase	SDS SCID	14,3%	5,6%	19,8%
Aragonés et al. (2005)	906	Incidental/doble fase	SDS SCID	16,8% *	–	–
Caballero et al. (2008)	8.687	Incidental/doble fase	GADS MINI	14%	–	–
Ferrer y Rodríguez (1999)	993	Estratificado/una fase	GHQ BDI	13,1%	8,6%	13,1%
Serrano et al. (2010)	3.851	Estratificado/una fase	SCID MINI	9,6%	5,4%	12,1%
King et al. (2008)	1.270	Incidental/una fase	CIDI	–	11,2%	18,4%
Estudios anteriores:	17.363	–	–	13,2% [†]	7,0% [†]	15,0% [†]

Nota. BDI = Inventario de Depresión de Beck; MINI = MINI Entrevista Neuropsiquiátrica Internacional; SDS = Escala Autoaplicada para la Depresión de Zung; SCID = Entrevista Clínica Estructurada para los Trastornos del Eje I del DSM-IV; GADS = Escala de Ansiedad y Depresión de Goldberg; GHQ = Cuestionario de Salud de Goldberg; CIDI = Entrevista Internacional Diagnóstica Compuesta.

* Datos conjuntos de episodio depresivo mayor y distimia. [†] Media ponderada por el número de pacientes de cada estudio.

1.2.2. Comorbilidad de los trastornos depresivos en atención primaria

En la práctica clínica, los cuadros afectivos suelen ir de la mano de otras entidades diagnósticas. A veces los trastornos depresivos y los trastornos que los acompañan son primarios, mientras que en otras ocasiones los trastornos depresivos aparecen de manera secundaria a una enfermedad física. Concretamente, la asociación entre trastornos de ansiedad y de depresión es prácticamente la norma más que una excepción. Existe una importante controversia sobre la consideración de ambos constructos como nosologías diagnósticas diferentes (Sanz, 1993). Como veremos en capítulos posteriores, instrumentos contrastados de evaluación de los trastornos depresivos presentan, con frecuencia, una validez discriminante moderada o incluso baja con la ansiedad. Demyttenaere et al. (2009) encuentran en su estudio que los médicos de AP tienen dificultad para discriminar entre ansiedad y depresión, tendiendo a diagnosticar más un cuadro que abarque ambas entidades. Las dificultades también aparecen con el diagnóstico diferencial de otros cuadros o con el diagnóstico comórbido de otros cuadros. Por enumerar ejemplos reseñables, existen dificultades para hacer el diagnóstico diferencial de ansiedad y depresión respecto a los trastornos adaptativos, en los cuales aparecen síntomas emocionales o comportamentales en respuesta a un estresor identificable dentro de los tres meses siguientes. Entre sus especificaciones aparecen los estados mixtos ansioso-depresivos donde como su propio nombre indica coexisten ambos tipos de síntomas. Frecuentemente entre las personas con rasgos marcados o incluso trastornos pertenecientes al cluster B (dramáticos, emotivos o inestables, DSM-IV-TR; 1996) aparecen

rasgos neurotiformes donde la confusión para el médico de AP entre los rasgos depresivos y ansiosos puede hacerse más patente.

Tal es la evidencia de la comorbilidad ansiedad-depresión, que el grupo de trabajo de los trastornos afectivos del DSM-V (APA; 2010, prevista su publicación para mayo del 2013), está considerando la inclusión de una dimensión de ansiedad dentro de los trastorno afectivos, dado que si bien es verdad que se pueden realizar ambos diagnósticos, la consideración de una misma entidad, permitiría según los autores, un mejor conocimiento de la entidad comórbida así como una planificación óptima del tratamiento desde el principio teniendo en consideración ambos aspectos. Actualmente en estudio (Apéndice B del DSM V: Criterios y ejes en estudio), los criterios propuestos serían: que el paciente tendría que tener 3 ó 4 síntomas de depresión mayor (debe incluirse tristeza y/o anhedonia) y acompañarse por síntomas de ansiedad (miedos irracionales, dificultad para relajarse, preocupaciones, anticipaciones acerca de que algo horrible ocurra...). Los síntomas deben estar presentes al menos desde hace dos semanas. No debe haber presente un diagnóstico de ansiedad o depresión según el DSM. Ambos deben estar concurriendo al mismo tiempo.

Quien sí parece hacerse eco desde hace tiempo de esta entidad común es la clasificación diagnóstica CIE-10 (OMS, 1993), pero la considera dentro de los trastornos de ansiedad. El diagnóstico del trastorno mixto ansioso-depresivo (F41.2) de la CIE-10 requiere que estén presentes síntomas de ansiedad y de depresión, pero ninguno de ellos predomine claramente ni tenga la intensidad suficiente como para justificar un diagnóstico por separado.

Parece entonces justificado proseguir con mayor investigación en el tema para poder aclarar la diferenciación clara o no de ambos cuadros clínicos.

En España, un estudio encuentra una elevada comorbilidad de los trastornos depresivos con otros trastornos mentales, lo que provoca en multitud de ocasiones que se evalúe y trate una enfermedad y no se tenga en consideración otras asociadas. De hecho, en el 46% de los pacientes con depresión mayor coexiste un trastorno mental más, dos en el 20% y tres trastornos mentales más en el 8%. El trastorno de ansiedad generalizada está presente en el 55% de los deprimidos, el trastorno de pánico en el 34%, la distimia en el 16% y el trastorno de somatización en el 7%. En los grupos de pacientes con enfermedades comórbidas la depresión fue más grave y con mayor impacto funcional (Aragonés et al., 2006).

Existe una abundante bibliografía en relación con la comorbilidad entre depresión y otros trastornos mentales en AP. Existe bastante acuerdo que son con los trastorno de ansiedad con los que más se relacionan, en particular con la fobia social y el trastorno por angustia (Serrano-Blanco et al 2008). Otros estudios encuentran la distimia (OR = 73,1), seguido de trastorno de angustia generalizada (OR = 67.5), posteriormente ataque de pánico (OR = 41,8) y finalmente planes, ideación suicida (OR = 7,5) (Gabilondo et al., 2010).

Como se ha comentado en multitud de ocasiones en el presente trabajo, las depresiones en AP suelen mostrarse bajo síntomas somáticos, pero también las enfermedades físicas suelen acarrear cuadros mentales. Así, depresión, ansiedad y deterioro cognitivo son las consecuencias más comunes de las enfermedades físicas. Por ejemplo, en un estudio internacional a gran escala (WHO, 2008), el 52% de las personas con cardiopatía mostraba

síntomas de depresión y aproximadamente un 30% tenía síntomas para un episodio depresivo mayor. Diabetes e hipertensión frecuentemente coexisten con depresión y demencia. Un diagnóstico de SIDA incrementa el riesgo de padecer una enfermedad mental (WHO, 2008). Los trastornos depresivos estaban asociados en mayor medida a dolor crónico, enfermedades gastrointestinales, alergias crónicas y en menor medida a limitaciones sensoriales y diabetes, sugiriéndose en dicho estudio que la comorbilidad física tiende a cronificar el curso de la depresión (Serrano-Blanco et al., 2008).

Al final, de una forma u otra, el médico de AP va a encontrarse de una manera muy continuada con los trastornos del estado de ánimo.

Mucho se ha hablado sobre las depresiones con amplio cortejo somático en AP. Caballero et al. (2007) analizan los síntomas físicos asociados a una muestra de depresivos en AP. Encuentran que la mayor frecuencia estaba en el área del dolor. Un 85,5% de las personas encuestadas tenía al menos un síntoma específico de dolor. Se seguía por síntomas cardiopulmonares (81%); gastrointestinales (69%); ginecológicos-sexuales (65%) y finalmente cuadros pseudoneurológicos-conversivos (64%). Teniendo en cuenta los síntomas individuales, la mayor prevalencia fue dolor de espalda (presente en el 69%), palpitaciones (57%), dolor en miembros (57%), dolor en articulaciones (49%), y flatulencia (48%).

En este mismo estudio se analizaron las atribuciones etiopatogénicas realizadas por los médicos de AP en relación con los síntomas físicos. Los datos son variables; los más asociados con depresión fueron los cardiopulmonares (80% de los pacientes con palpitaciones, 71% de los pacientes con dolor torácico y 61% de los pacientes con disnea), los

gastrointestinales fueron atribuidos a génesis depresivas en la mitad de los casos (68% de los pacientes con náusea, 55% de los pacientes con flatulencia y 54% de los pacientes con abdominalgias). Los síntomas que cursaban con dolor eran atribuidos en aproximadamente uno de cada cuatro (24% pacientes con dolor de espalda, 24% de los pacientes con dolor en miembros y 16% de los pacientes con dolor articulario). La atribución de los síntomas conversivos y ginecológicos fueron más variables (Caballero et al., 2007).

García-Campayo et al. (2008) encontraron una relación significativa positiva entre el número de síntomas físicos, la gravedad de la depresión y los recursos de salud mental utilizados. A su vez encontraron una relación negativa entre el número de síntomas físicos y la calidad de vida. Aproximadamente dos tercios de los pacientes con depresión en AP presentaban síntomas somáticos.

Agüera et al. (2010) encontraron una prevalencia del 56% de episodio depresivo mayor en personas que consultaban por quejas dolor sin explicación física. Cabe resaltar la prevalencia también significativa de trastorno depresivo menor (18%) y de distimia (17%) en estas personas que sufrían dolor.

El conocimiento por parte de los médicos de AP de este tipo de investigaciones permite la mejora de la identificación y la consideración de la depresión en su más amplio espectro y complejidad.

1.3. Importancia sociosanitaria de los trastornos depresivos

1.3.1. Repercusiones de los trastornos depresivos en la sociedad

La depresión es una de los trastornos mentales más frecuentes en nuestra sociedad. Las repercusiones personales, familiares, laborales y socioeconómicas de la depresión son estremecedores. La calidad de vida del individuo y de su familia se ve afectada habitualmente de forma más intensa que en otras enfermedades. En los trastornos depresivos se produce una importante incapacidad y sufrimiento de las personas que lo padecen. Los trastornos del estado de ánimo se asocian con el suicidio en un alto porcentaje de casos, de tal forma que se calcula que hasta un 15-20% de los casos con depresión mayor tienen un alto riesgo de finalizar en suicidio (Gabarrón et al., 2002).

En el estudio del *Global Burden of Disease* (Murray y López, 1997), se estima que la depresión unipolar será la segunda causa de muerte mundial en el 2020, tan solo precedido por la enfermedad isquémica coronaria. En otro estudio más reciente de el *Global Burden of Disease* del 2001, se encuentra que la depresión unipolar se situaba en la tercera posición en incapacidad funcional (medida a través de años de vida perdidos) en los países de mayores recursos y la séptima en los países de ingresos medios y bajos. (Citado en López et al., 2006).

Está previsto que para el 2030 la depresión unipolar pasará a ser la segunda causa de carga social (medido a través de incapacidad funcional y pérdida de días trabajados) a nivel global, exclusivamente precedido por el sida

y en tercer grado estarían las isquemias cardíacas (Mathers y Loncar, 2006). Actualmente parece que la depresión es la primera causa de jubilación precoz o absentismo laboral en muchos países europeos (Curran et al., 2007).

Greenberg y Kessler (2003) apuntan que el incremento de la carga económica de la depresión se debe a la pérdida de la funcionalidad, productividad y a la mayor utilización de servicios médicos, y a que éstos exceden los recursos dedicados al tratamiento. En España la prevalencia de la depresión en el ámbito laboral representa aproximadamente un 10%, siéndole atribuible un 5% de las bajas laborales. Constituye una de las causas más frecuentes de absentismo laboral de larga duración, con una media de 6 meses de baja por episodio (Chamorro, 2004).

Un reciente estudio encuentra que además de los gastos directos acarreados por la depresión (consulta médica, hospitalización, costes farmacéuticos...), la mayoría de los estudios no consideran el gasto indirecto (pérdida de productividad, uso de recursos de la comunidad, cuidadores informales...), suponiendo este gasto indirecto aproximadamente la mitad del monto total. Además, concluyen los autores, que los costes habría que considerarlos según la fase en la que se encuentre la persona, de tal manera que habría diferencias significativas si se encuentra en fase aguda o en remisión. Realizan a partir de los datos del estudio LIDO (2002) una estimación del gasto en 2009 en España variando el coste de 334€, 874€ a 1335€ por año según se encuentre en fase de remisión, remisión parcial o aguda respectivamente (Karapampa, Borgstrom y Jonsson, 2011).

Datos como los anteriormente expuestos indican la importancia de los problemas que acarrearán los trastornos depresivos no sólo a nivel del individuo,

sino a nivel comunitario, como un problema social con importantes repercusiones socio-económicas.

La elevada prevalencia de los trastornos del estado de ánimo a nivel internacional justifica el hecho de sopesar y poner en marcha políticas dirigidas a la mejora de la atención en salud mental y primaria. De hecho, la OMS en su último informe refiere que “determinadas habilidades y competencias son necesarias para realizar una evaluación, diagnóstico, tratamiento, apoyo y derivación efectivos en personas con trastorno mental. Es necesario que las trabajadoras de AP estén adecuadamente preparados y apoyen el trabajo de salud mental” (*World Health Organization*; WHO, 2008, p. 42).

1.3.2. Repercusiones de los trastornos depresivos a nivel de salud pública

Actualmente la AP constituye el primer nivel asistencial del sistema sanitario en España. Desde el proceso de la Reforma Psiquiátrica en los años 80 del siglo pasado, es la vía de entrada para los servicios de atención especializada, especialmente a los centros de salud mental que se erigen como eje interconector de los diferentes dispositivos de 2º nivel o especializada. Es en ese primer nivel de AP donde se manejan fundamentalmente los trastornos mentales más frecuentes en la población, como son los trastornos depresivos. De hecho, la mayoría de los pacientes que sufren depresión son atendidos sustancialmente en este nivel asistencial (Wittchen, Holsboer y Jacobi, 2001; Cuijpers, Straten, Schaik y Andersson, 2009). En un estudio del 2007 apuntaron que las enfermedades mentales eran la cuarta causa de visita al

médico de AP, ocupando concretamente la depresión y la ansiedad el sexto puesto (Binns et al., 2007).

En el informe de la OMS del 2008 refieren que los principales trastornos mentales en AP son la depresión (variando desde un 5% hasta un 20%), el trastorno de ansiedad generalizada (4-15%) consumo y dependencia de alcohol (5% -15%) y trastornos por somatización (0,5-11%) (WHO, 2008). En España, Arbesu Prieto et al. (2006) encuentran que los trastornos mentales más frecuentes en AP son los de corte ansioso-depresivo, a saber: trastorno de ansiedad (13%), trastorno de ansiedad-depresión (13%), depresión (10%), comorbilidad ansiedad-depresión (19%) y trastorno adaptativo (3%).

El uso de recursos asistenciales del paciente depresivo, una vez que se introduce en el sistema sanitario, se calcula que es tres veces superior al de la media del resto de pacientes, con la inevitable pérdida de productividad en el ámbito laboral.

Desde los últimos años, se hace patente la necesidad de considerar la depresión como una enfermedad crónica. Dada la alta frecuencia de periodos en los que se alterna la recaída, periodos de remisión parcial/total y/o recurrencia, parece fundamental la concepción de la depresión como una enfermedad mental con unas tasas de morbilidad muy relevantes a nivel internacional.

1.4. Dificultades para el diagnóstico de los trastornos depresivos en atención primaria

Como se ha constatado en epígrafes anteriores, la carga asistencial de pacientes que padecen trastornos depresivos es muy elevada. Las dificultades en la precisión diagnóstica acarrearán consecuencias de diversa índole. Puede que la persona afectada de depresión pase inadvertida para su médico de cabecera, cronificándose el cuadro. A veces se considera depresión cuadros subclínicos con la consiguiente derivación a salud mental, prescripción de tratamiento hasta ser visto por el especialista, confusión en expectativas de mejoría..., con la consecuente sobrecarga de servicios, tanto a nivel de salud mental como de la propia AP (p. ej., el paciente tendrá que acudir a AP para la prescripción de medicación, revisiones de incapacidad laboral transitoria, etc.). En ocasiones el médico de AP se sobrecarga con personas con trastornos depresivos graves sin conocer que lo padecen y sin derivarlos al especialista en salud mental.

Se estima que solamente 1/3 de los pacientes depresivos acuden inicialmente a un especialista de salud mental. Por tanto, el médico de AP es actualmente una figura clave en la detección de los trastornos depresivos (Gabarrón et al., 2002).

La detección de los trastornos depresivos por parte del médico de AP tiene una relevancia fundamental en la prevención de suicidios asociados al trastorno, en la mejora de la calidad de vida del paciente y en la reducción de costes sanitarios y sociales, dada su hiperfrecuentación de los servicios médicos y la incapacitación que conllevan (son la primera causa de años

vividos con incapacidad; Murray 2001). En esta misma línea, Lake (2008) considera que la capacidad del médico de AP para detectar trastornos depresivos es de importancia fundamental en la prevención de suicidios relacionados con la depresión así como para aumentar la calidad de vida de los pacientes.

Los grupos con peor porcentaje de detección son las personas viudas, jubiladas, hiperfrecuentadores y las mujeres (Gabarrón et al., 2002). Los que más fácilmente se reconocen son los casos más graves y los más incapacitados.

En el estudio de Cepoiu et al. (2008) se demuestra que los médicos de AP no tienen peor capacidad de detección que otros médicos no psiquiatras, por lo que la posibilidad de mejora en la identificación, no sólo ayudaría al médico de AP, sino también a los médicos de especialidades ajenas a la salud mental.

1.4.1. Infradiagnóstico y sobrediagnóstico de los trastornos depresivos en atención primaria

Las dificultades en la precisión diagnóstica en depresión en el marco de AP es un hecho bastante refrendado. Se calcula que en AP se produce un infradiagnóstico de un 1/3-1/2 aproximadamente de los casos de depresión (Gabarrón et al., 2002). Un estudio posterior de estos mismo autores (Gabarrón et al., 2004) encontraron que permanece infradiagnosticado un 44,3% de las personas con probable trastorno depresivo (sobre todo mujeres, viudos,

jubilados, los que han presentado acontecimientos vitales estresantes y los individuos más frecuentadores de la consulta).

La OMS encuentra una gran variabilidad en el porcentaje de reconocimiento de la depresión por parte de los médicos de AP, variando desde un 14% en Japón hasta un 70% en Verona (Italia) o en Manchester (Reino Unido), de manera que, como media, el médico de AP detectaría un 39,1% de las personas con depresión (WHO, 2008).

Mitchell et al. (2009) en su meta-análisis de 41 estudios encuentran una identificación correcta del 47% (95% IC = 42-53%). Es decir, el médico de AP identificaría aproximadamente a la mitad de los verdaderos positivos. Cabe resaltar la importante diferencia en sensibilidad si se analizan los estudios por separado variando desde un 7% hasta un 79%. Cuando el estudio se hace retrospectivo, analizando las historias clínicas, la identificación correcta por parte del médico de AP se reduce a un 34% (95% IC = 22-46%). La especificidad mejora en los médicos de AP ya que son capaces de identificar correctamente a un 81% de las personas no deprimidas. En este estudio de Mitchell et al. (2009) también encontraron que con la reevaluación, la detección mejoraba en un 35%, es decir, considerando una sensibilidad del 50% y una especificidad del 80%, si cada caso donde cabe sospecha es reevaluado, entonces la precisión total de los clínicos aumentaría hasta un 89%.

En los últimos años se ha tratado de estudiar la concordancia entre el diagnóstico de depresión formulado por el médico de AP y diferentes autoinformes. De este modo, se encontró que de los pacientes que habían resultado positivos a través del cuestionario de cribado *Patient Health Questionnaire* (PHQ-9), sólo el 31% había sido identificado como depresivo por

el médico de AP. El índice de acuerdo encontrado entre el médico de primaria y el PHQ-9 medido a través del índice kappa fue bajo ($\kappa = 0,001-0,101$; Ani et al., 2008).

En la misma línea, un estudio reciente (McGrady, Lynch, Nagel y Tamburrino, 2010) estudió la coherencia entre escalas automedidas: PRIME MD (Spitzer et al., 1994), *Beck Depression Inventory* (Beck, 1961) y *Eysenck Personality Inventory* (EPI; Eysenck and Eysenck, 1994) con las anotaciones procedentes de las historias clínicas de los pacientes. Encontraron un índice kappa de concordancia de 0,077, lo que es considerado como un resultado bajo. La correlación entre el PRIME-MD y la subescala de neuroticismo del EPI fue de 0,55, lo cual, es considerada como adecuada. Aproximadamente la mitad de los pacientes identificados como positivos por los autoinformes no eran identificados como tales por los médicos de AP.

Cabe tener en cuenta la importancia de un diagnóstico preciso no sólo relativo al infradiagnóstico, sino también considerar que el hecho de tener falsos positivos conlleva en muchas ocasiones la estigmatización del paciente y la asunción del rol de enfermo, y, además, la prescripción de fármacos antidepresivos innecesarios.

Aragonés, Piñol y Labad (2006) encuentran en su estudio que un 26% de personas fueron diagnosticadas como depresivas cuando realmente no cumplían criterios formales para la depresión. El estudio europeo ESEMED encuentra que un 13% de las personas que son tratadas presentan síntomas de depresión sin presentar cuadro clínico. En el estudio americano NCS-R el

dato es más alarmante, un 40% de las personas que reciben tratamiento para depresión no presentan un síndrome concurrente (Mitchell et al., 2009).

1.4.2 Factores asociados a las dificultades diagnósticas de la depresión

Mitchell, Vaze y Rao (2009) detallan algunos de los factores asociados a la baja detección de la depresión por parte del médico de AP, entre los cuales se encontrarían:

1. Prevalencia de depresión en AP es baja en relación a especializada.
2. No expresión explícita de quejas emocionales. En los casos en los que se verbalizan las quejas psicológicas mejora la detección.
3. Dificultad en clasificar la complejidad patognomónica de la depresión en categorías diagnósticas, especialmente en los casos leves o subclínicos.
4. Dificultad percibida por el paciente para la exposición de los síntomas de depresión teniendo en cuenta el poco tiempo con que cuentan los médicos de AP por consulta.
5. Que el médico de AP priorice y se interese por aspectos socioeconómicos del paciente (laborales, económicos, vivienda...).
6. Otros aspectos que tienen que ver con la relación terapéutica, presentar síntomas psicológicos, experiencia clínica del médico, acceso, contacto con el paciente, etc.

Linarez (2004) encuentra en su estudio varias dificultades y limitaciones para el diagnóstico de la depresión por los médicos de AP, entre ellos, los relacionados con la escasa formación psiquiátrica de los médicos de AP, sus

condiciones precarias de trabajo y la tendencia a orientar su atención hacia los trastornos físicos más que hacia los psicológicos.

Según otros estudios, el principal problema diagnóstico en el momento actual es el enmascaramiento con otra sintomatología (76%) así como la falta de tiempo en la consulta (67%) (Martín-Águeda et al., 2005, 2007). García-Campayo et al. (2008) consideran que dos tercios de los pacientes con depresión en AP presentan síntomas somáticos, lo cual, complica el diagnóstico.

Gabarrón et al. (2002) encuentran diversas razones para explicar el infradiagnóstico de la depresión por parte de los médicos de AP:

1. Pluripatología orgánica que en muchas ocasiones se solapa con síntomas depresivos.

2. Características de los pacientes depresivos que acuden a las consultas de AP: de menor gravedad y mayor evolución respecto a los que acuden a especializada (casos recurrentes, con una clínica más florida, que han sido precozmente derivados a salud mental y cumplen mejor los criterios diagnósticos psiquiátricos).

3. Discontinuidad en el seguimiento por parte del médico de AP. Se mejoraría la detección a mayor conocimiento y familiaridad del médico de AP con la historia psiquiátrica del paciente.

Posteriormente, Chinchilla (2008) realiza una relación de dificultades de los médicos de AP en el diagnóstico de los trastornos depresivos, entre los que destacarían:

- Formación de pregrado más academicista que auténticamente clínica.

- Formación continuada no siempre adecuada o suficiente.
- Posible confusión entre lo sintomático y lo sindrómico depresivo.
- Dificultades diagnósticas (enmascaramientos somáticos).
- Dificultades entre lo categorial y lo dimensional depresivo.
- Alta frecuencia de comorbilidad entre depresión y otras patologías psiquiátricas o médicas (cáncer, alcohol, diabetes, etc.).
- Alta frecuencia de síntomas ansiosos y depresivos

El Prof. Chichilla también achaca las dificultades del médico de AP en la detección de la depresión al triste y cada vez mayor descrédito de la fenomenología por los especialistas, en pro de las entrevistas con cuestionarios, lo que, según él, generaría una menor capacitación en semiología psicopatológica. Sin embargo, esta opinión no está basada en ningún dato y es totalmente discutible. De hecho, como se verá en el capítulo siguiente, existen muchos estudios que sugieren que la utilización de cuestionarios de cribado por parte del médico de AP aumenta la probabilidad de que éstos detecten trastornos depresivos entre sus pacientes.

Por su parte, la OMS (2008) señala dificultades diferenciales asociadas a los pacientes, los profesionales sanitarios y el sistema sanitario, todas las cuales inciden en que los trastornos depresivos pasen desapercibidos para el médico de AP.

En cuanto a los factores asociados a los pacientes, la OMS señala:

- Focalización en síntomas físicos en detrimento de los psicológicos.

- Considerar que no se necesita ayuda profesional para los problemas psicológicos.
- Miedo a la hospitalización en salud mental, o al estigma o vergüenza por la utilización de los servicios psiquiátricos.

En relación con los factores asociados a los profesionales, la OMS señala:

- Formación precaria en salud mental.
- Incomprensión sobre la naturaleza de los trastornos mentales, y prejuicios o tiempo insuficiente para evaluar los trastornos mentales.
- Presentación somática de cuadros mentales.
- Poco interés en los trastornos mentales y fácil derivación de los pacientes a salud mental.
- Padecer desgaste profesional.
- Sentir incapacidad ante la problemática social o el estigma.

Finalmente, la OMS destaca entre los factores asociados al sistema sanitario:

- Financiación y recursos humanos insuficientes, especialmente en países con bajos recursos.

En conclusión, a partir de los estudios sobre el tema parece evidente que una mejor coordinación y comunicación entre especialistas de salud mental y AP mejoraría la idoneidad de las detecciones de trastornos depresivos y de las derivaciones. Además no se puede olvidar que la rotación de los médicos de AP en servicios de psiquiatría es de apenas tres meses, con lo que la

formación del futuro médico de AP depende más bien de su voluntad, actitud y motivación por la salud mental que de una docencia adecuada. La formación continuada desde los servicios de salud mental debería estar cada día más presente. Actualmente se está gestando en muchas comunidades españolas la figura del psicólogo clínico/psiquiatra consultor en AP con el fin de mejorar la formación y comunicación entre ambos. De esta forma se potencia la capacitación del médico de AP en casos de trastornos afectivos leves, subclínicos, o que se puedan manejar con herramientas psicofarmacológicas sin necesidad de revisión por los servicios de psiquiatría.

El uso de instrumentos de evaluación como cuestionarios, inventarios o escalas, lejos de denotar un detrimento en el conocimiento de la semiología, psicopatología o nosología de los trastornos mentales, permiten al médico de AP, de una manera eficaz, explorar sintomatología de forma más exhaustiva. Dotar a los médicos de AP con instrumentos válidos y fiables facilitará su labor clínica. No se puede olvidar que la decisión final sobre el diagnóstico vendrá con otras pruebas complementarias y el criterio clínico del profesional sanitario.

1.5. Métodos de detección de la depresión por parte de los médicos de atención primaria

Los instrumentos que de forma mayoritaria suelen utilizar los médicos de AP para diagnosticar a los pacientes con depresión son la entrevista (90%) y la valoración de la sintomatología (89%) (Martín-Águeda et al., 2007). A pesar de las bondades de estos instrumentos, también presentan serios problemas y, de hecho, los médicos de AP suelen identificar correctamente algunos síntomas

depresivos como la tristeza o la apatía, pero les cuesta identificar otros síntomas depresivos tales como los sentimientos de inutilidad o los problemas de concentración (véase la Tabla 1.4 en la que se presenta la sintomatología depresiva que los médicos de AP suelen identificar con mayor probabilidad).

Dadas estas dificultades, parece relevante facilitar un método sencillo y sistemático de evaluación que permita al médico de AP considerar los criterios diagnósticos, así como para poder diferenciar los trastornos depresivos de otros que puedan resultar similares como por ejemplo los trastornos adaptativos o el malestar propio de la vida cotidiana.

Estos aspectos se pueden ver mejorados con un instrumento de detección, cribado o despistaje (*screening*) que optimice la información sobre el posible trastorno que puede obtener el médico de AP.

Tabla 1.6. Sintomatología depresiva identificada correctamente por el médico de atención primaria (Martín-Águeda et al., 2005)

SÍNTOMAS	% de IDENTIFICACIÓN por el médico de atención primaria
Estado de ánimo deprimido (tristeza, labilidad emocional, llanto)	75,6%
Disminución de la vitalidad (apatía, anergia, astenia, inhibición)	48,7%
Anhedonia	36,1%
Somatización	24,8%
Ansiedad	15,1%
Trastornos del apetito	13,4%
Ideas de suicidio	7,1%
Sentimientos de inutilidad	4,6%
Falta de concentración	3,4%

En resumen, y siguiendo las consideraciones de Caballero y García-Parajuá (2004), es necesario la ampliación de los médicos de AP como agentes imprescindibles en el diagnóstico y tratamiento de la depresión, respecto a la consideración de los especialistas en salud mental como únicos facultativos competentes en el abordaje de los trastornos del estado de ánimo. De tal modo que, en el plano clínico, la elevada prevalencia de la depresión y su influencia negativa en el curso de otras enfermedades como factor de riesgo o de pronóstico negativo, la relativa facilidad de su diagnóstico y la actual disponibilidad de tratamientos eficaces fáciles de aplicar, fundamentan la necesidad de que todo clínico generalista esté capacitado para identificar y tratar un buen número de depresiones no complicadas.

En el plano asistencial, cualquier planteamiento terapéutico de un problema de salud de la magnitud y características de la depresión, hace imprescindible la participación del sistema primario de salud, que por su accesibilidad, capacidad de seguimiento continuado, atención global y coordinación, parece el marco idóneo de actuación desde una perspectiva de salud pública.

Capítulo 2:

Uso de instrumentos de cribado o despistaje (*screening*) de los trastornos depresivos

El hombre es la medida de todas las cosas;

para las que son, medida de su ser;

para las que no son, medida de su no ser.

Protágoras

2.1. Conceptualización y uso de los instrumentos de cribado o despistaje (*screening*)

La evaluación puede definirse como el proceso de aplicación de un método sistematizado para describir fenómenos u objetos (Carulla, López Alberca y Sáiz, 2000). Habitualmente se diferencia entre medición (referido al proceso de recogida de la información) y evaluación (integración y valoración de la información recogida) (Cronbach, 1990). De manera popular, se entienden ambos conceptos como “evaluación”. El término psicodiagnóstico desde un punto de vista de diagnóstico psicopatológico, tiene como fin el establecer la clasificación del sujeto objeto de estudio con fines clínicos (Fernández-Ballesteros, 2010). Uno de los métodos más comúnmente usados para el psicodiagnóstico y la medición de atributos son los tests, cuestionarios o escalas psicopatológicas. Aunque existen diferencias entre estos últimos tres tipos de instrumentos, aquí se consideraran como sinónimos para referirse a procedimientos de medida psicopatológica cuyo material, forma de administración, puntuación, corrección y valoración son estándar y de los que se derivan puntuaciones normativas (Fernández-Ballesteros, 2010).

Los tests, cuestionarios o escalas psicopatológicas pueden utilizarse para muchos objetivos clínicos como, por ejemplo, el diagnóstico de trastornos, la evaluación sintomatológica, la evaluación del progreso terapéutico individual, la valoración de programas de tratamientos, etc. Entre estos objetivos está la detección, cribado o despistaje de trastornos, para los cuales se han construido un buen número de tests, cuestionarios y escalas con características específicas.

A la hora de decidir qué instrumento de cribado se ha de elegir para la práctica profesional o para la investigación, existen algunos aspectos que son relevantes a tener en consideración para fundamentar la decisión de elegir uno u otro test de cribado. Estos criterios (Muñoz, 2000) son importantes no sólo cuando se va a utilizar el instrumento como detector de casos, sino también en otras áreas (p. ej., investigación, comparación entre estudios, etc.):

- ¿Ha sido construido para los objetivos para los que se pretende utilizar?
- Si ha sido traducido, ¿la adaptación es adecuada para el nuevo contexto cultural?
- ¿La validación ha sido realizada de manera adecuada?
- ¿Conserva la capacidad de predicción de la versión original?
- ¿Se ha llevado a cabo su validación en una población semejante a la que se va a estudiar?
- ¿El punto de corte recomendado se adapta a los objetivos o tendría que modificarse?

Dado el gran número de instrumentos de que se dispone actualmente, parece necesario valorar con cuidado las garantías psicométricas que presenta cada uno de ellos, ya que, además, los resultados que se deriven del instrumento van a traer repercusiones importantes a todos los niveles y en sus diferentes usos.

2.1.1. Definición de cribado o despistaje (*screening*)

Cribado se refiere a la acción y efecto de cribar o seleccionar rigurosamente (Diccionario de la Lengua Española de la Real Academia Española, vigésimo segunda edición: <http://buscon.rae.es/drael/>). “Cribado” es un término genérico que se utiliza en ámbitos clínicos como traducción de la palabra inglesa *screening*. Esta última palabra también se puede traducir en español como “despistaje”, ya que, aunque el Diccionario de la Lengua Española define despistaje de forma muy restrictiva como “examen médico preventivo para detectar cáncer o enfermedades venéreas” (Diccionario de la Lengua Española de la Real Academia Española, vigésimo segunda edición: <http://buscon.rae.es/drael/>), el uso común de la misma es aplicarla a cualquier examen clínico para detectar patologías.

La génesis y posterior desarrollo de los instrumentos de cribado en psicopatología se llevan a cabo a partir de la Segunda Guerra Mundial, cuando debido a las necesidades del ejército americano para medir determinadas características clínicas de manera rápida, precisa y colectiva, se comienzan a utilizar los primeros test de cribado para la enfermedad psiquiátrica (Stouffer et al., 1950).

Las pruebas o test de cribado en el ámbito sanitario tienen como objetivo establecer de forma estandarizada una diferenciación probabilística entre población sana y enferma (Muñoz, 2000). La OMS publicó en 1968 un libro que versaba sobre los principios y la práctica del cribado en la enfermedad y que tuvo una gran repercusión en el desarrollo de este tipo de pruebas o tests. En esos momentos el cribado estaba cobrando cada vez mayor relevancia y

surgen de manera paralela discrepancias respecto a sus características y usos. En su publicación, la OMS adoptó la definición de cribado propuesta por la Comisión de enfermedades crónicas de Estados Unidos de América (*Comission of Chronic Illness; CCI*) en la que se define el *screening* como “la presunta identificación de enfermedades o defectos sin identificar mediante la aplicación de test, exámenes u otros procedimientos que pueden ser aplicados rápidamente. Los test de cribado separan aparentemente bien, personas que probablemente tienen una enfermedad de aquellas que probablemente no”. En resumen, la OMS considera que el objeto del cribado de enfermedades es “descubrir, en aquellos que aparentemente están bien, quienes en realidad están sufriendo una enfermedad” (WHO, 1968).

La OMS refería en su libro sobre el cribado (WHO, 1968) una serie de principios a considerar en relación con la práctica clínica del despistaje, especialmente cuando se considera en términos de salud pública:

- La condición objeto de estudio debe ser un problema de salud importante.
- Una vez la enfermedad ha sido reconocida, debe haber un tratamiento validado para el paciente.
- Deben estar disponibles acceso y facilidades para el diagnóstico y tratamiento.
- Debe encontrarse la sintomatología en estado latente reconocible o en un estadio precoz.
- Debe haber un test o examen apropiado.
- El test debe ser aceptable para la población.

- La evolución natural, incluyendo el desarrollo de enfermedad latente a declarada, debe ser adecuadamente comprendida.
- Debe haber un acuerdo en el protocolo a seguir sobre cómo tratar a los pacientes.
- El coste del cribado (incluyendo diagnóstico y tratamiento de los pacientes diagnosticados), debe ser económicamente equilibrado en relación con los posibles gastos médicos en general.
- El estudio de caso debe comprenderse dentro de un proceso continuo y no como un proyecto aislado.

Una definición muy influyente también sobre el cribado es la que ofrece el Comité Nacional de Cribado del Reino Unido (*National Screening Committee UK*). Este comité, creado en 1997, se encarga del asesoramiento en esa materia en los cuatro países que conforman el Reino Unido y tiene como objetivo identificar programas de cribado que hayan evidenciado tener mayores beneficios que daños. Según dicho comité, el cribado es el proceso de identificación de personas, aparentemente sanas, que pueden tener un alto riesgo de enfermedad (Gray, 2004). Además, según este comité, el cribado “puede ofrecer información, así como pruebas diagnósticas y tratamiento adecuado con el fin de reducir el riesgo de padecer alguna complicación emergente secundaria a la enfermedad o condición” (citado en su página web: <http://www.screening.nhs.uk/screening>). Por tanto, la idea que en los últimos años se trata de transmitir es ver el cribado como un programa comprehensivo dentro del control de la enfermedad, no sólo como la administración de un test (Gray, 2004).

Los requisitos de cribado que marca el Comité Nacional de Cribado del Reino Unido serían:

- La condición debe ser suficientemente común para merecer un cribado. Esto no significa necesariamente ser común en la población general, a no ser que se trate de un cribado universal. Significa común en el grupo diana para el cribado.
- Debe existir una intervención efectiva para la enfermedad sujeto de cribado.
- El cribado debe realizarse en los estadios precoces, cuando la intervención es más efectiva.
- El método de detección debe tener una alta especificidad y una muy alta sensibilidad, aunque se considera que este aspecto es difícil de evaluar cuando se trata de una herramienta de cribado de depresión.
- El test debe ser relativamente barato o de lo contrario, el coste por caso detectado será prohibitivo.
- El test debe ser seguro, fácil, estar validado para la muestra objeto de estudio, ser preciso y aceptado por el paciente.
- Debe ser conocida la distribución de las puntuaciones en la población diana, así como definido un punto de corte adecuado.
- En los casos positivos, debe haber un protocolo para proseguir con la evaluación y diagnóstico, así como ofrecer las múltiples alternativas posibles para ese individuo.

En la misma línea y esta vez en España, Muñoz (2000) resume las características de un test de cribado:

- Facilidad de administración.
- Bajo coste.
- Aceptable por la población.
- Contenidos no contaminados por variables de confusión.
- Sensibilidad suficiente: más de 0,50 para los casos.
- Correlación elevada con el criterio externo.
- Alto poder discriminativo entre sanos y enfermos

También existen categorizaciones de cribados en función de la longitud, entre las que destacarían:

- Ultracortos: Son los formados entre uno y cuatro ítems, necesitando menos de dos minutos para completarlo.
- Cortos: Son los compuestos entre cinco y quince ítems, necesitando entre dos y cinco minutos para registrarlos.
- Típico o habitual: Son los compuestos por quince o más ítems, necesitando más de cinco minutos para rellenarlos.

Concretamente en AP, un buen instrumento de cribado es un instrumento de evaluación que se caracteriza por ser breve para poder ser aplicado en las visitas médicas sin consumir excesivo tiempo, por ser sencillo y práctico de administrar, por requerir un tiempo mínimo de aprendizaje y entrenamiento, por ser aplicable a todos los pacientes, y por ser exhaustivo o completo para detectar el cuadro que se quiere evaluar. Además, los

procedimientos de cribado deben ser altamente sensibles y específicos, susceptibles de ser reproducidos por otro profesional, y ofrecer una valoración de riesgos, molestias y costes.

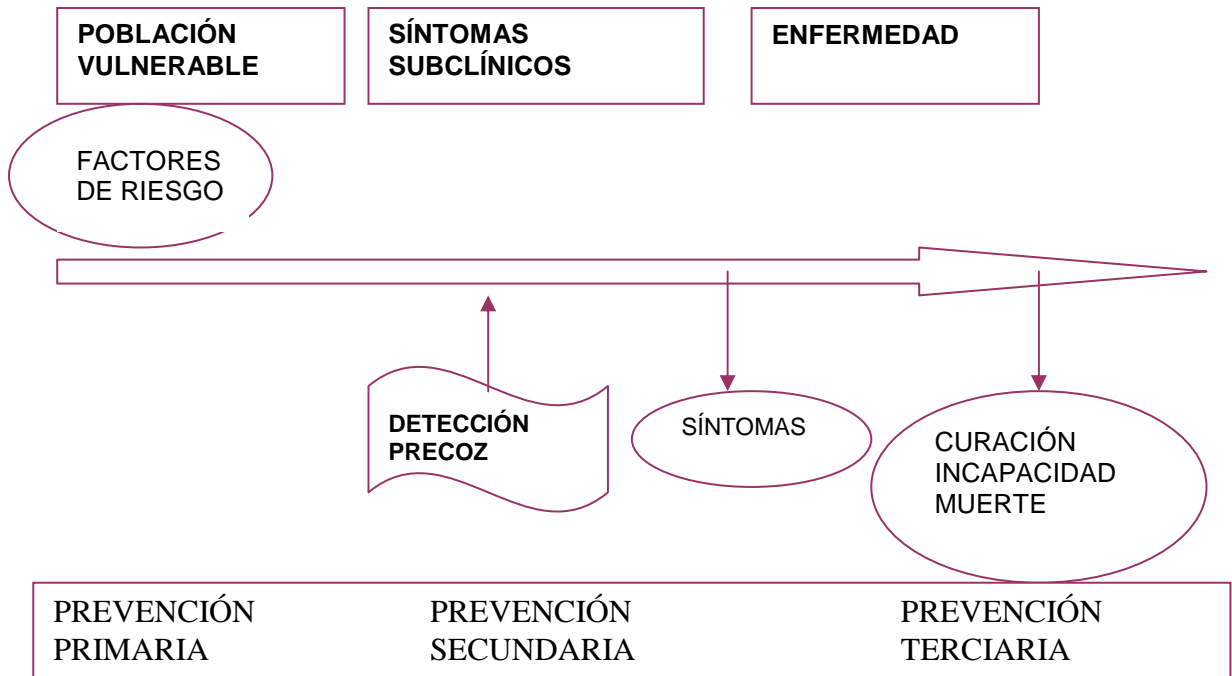
2.1.2. Usos habituales del cribado

El uso del cribado puede hacerse a diferentes niveles, desde poblacionales (para conocer datos epidemiológicos de una patología) hasta individuales para identificar probables casos clínicos. Para realizar un diagnóstico será necesario realizar estudios complementarios que confirmen los datos iniciales.

La utilización del cribado en AP respondería a un nivel de prevención secundaria, es decir, en el caso de la depresión, por ejemplo, el médico tiene sospechas clínicas de que una persona puede estar padeciendo un trastorno depresivo y necesitaría un primer método rápido y fiable para confirmar sus sospechas. Si el resultado del cribado fuese positivo, entonces el médico debería realizar exámenes complementarios para realizar el diagnóstico (véase la Figura 2.1).

Por lo tanto, uno de los objetivos principales del cribado en atención primaria sería la detección precoz de la enfermedad. De esta manera, se proveería de tratamiento a personas que padecen la enfermedad y no se les había detectado previamente y, por otro lado, se evitarían potenciales daños secundarios a un diagnóstico y tratamiento innecesarios.

Figura 2.1 Diferentes niveles de prevención de la enfermedad



La OMS diferencia diferentes tipos y usos de cribado (WHO, 1968):

- Cribado masivo:

Cribados realizados a gran escala en toda la población, es decir, serían aquellos procesos de despistaje donde no se realiza una selección de grupos previa.

- Cribado selectivo:

Se seleccionan grupos de riesgo de la población. Puede ser realizado a gran escala y entonces, puede considerarse una forma de *screening* poblacional.

○ Cribado múltiple o multifásico:

Consiste en la aplicación de dos o más tests de cribado combinados a un gran número de personas. Se realiza para economizar en tiempo y esfuerzo humano.

○ Vigilancia (*surveillance*):

Se tiende a entender el cribado de manera transversal, como un proceso a corto plazo en población de riesgo. Sin embargo, el cribado de vigilancia considera el despistaje como un proceso de vigilancia a largo plazo sobre la salud pública o individual, y, por tanto, requiere la realización de cribados repetidos a determinados intervalos de tiempo durante en un proceso a largo plazo.

○ Estudios de caso:

El objeto principal de este tipo de cribado es detectar una enfermedad y proponer a la persona un tratamiento en caso de ser necesario.

○ Estudios epidemiológicos:

Estudios con el fin de conocer la prevalencia, incidencia, evolución de la variable o variables objeto de estudio. Los estudios de caso se podrían considerar como un subproducto de éstos.

○ Detección precoz:

Se consideraría este significado de cribado cuando, en cualquier tipo de despistaje, el objeto de análisis es la detección precoz de la enfermedad.

2.2. Población diana susceptible del cribado de los trastornos depresivos

En prácticamente todas las guías clínicas sobre el uso del cribado en salud aparece la importancia del despistaje en población de riesgo, en este caso, aquellos subgrupos que por sus características se situarían con una mayor probabilidad de sufrir un trastorno depresivo.

En la comunidad catalana, en el plan de salud de su Departamento de Sanidad (Gencat/salut, 2007), se ha elaborado un documento que recoge la población de riesgo susceptible de cribado en depresión en función de distintos antecedentes personales y factores psicosociales.

En relación con antecedentes personales:

- Historia anterior de trastornos depresivos
- Antecedentes familiares de trastornos mentales graves (esquizofrenia, depresión mayor, trastorno bipolar)
- Antecedentes familiares o personales de intento de suicidio
- Enfermedad física crónica o severa
- Pacientes con manifestaciones psicopatológicas.
- Pacientes con síntomas somáticos inespecíficos (mareo, cefaleas, parestesias, dolor torácico)
- Pacientes consumidores habituales de psicofármacos o sustancias psicoactivas
- Pacientes pluriconsultantes.

En relación con los factores psicosociales:

- Niñez: pérdida o separación de los padres, pérdida de contacto con el hogar
- Adolescencia: separación de los padres, del hogar y de la escuela, fracaso escolar
- Adultos jóvenes: ruptura matrimonial, embarazo, nacimiento de un hijo disminuido, pérdida de trabajo, pérdida de progenitor, emigración y carencia de recursos sociales
- Adultos y ancianos: jubilación, pérdida de funciones físicas, luto, pérdida de un familiar, enfermedad en la familia.
- Responsabilidades de cuidar de otros (p. ej., cuidado de un familiar anciano)
- Otras transiciones psicosociales.

La guía NICE (2009) sugiere el uso de cribado en aquellos individuos con una historia pasada de depresión, con enfermedad física significativa (especialmente si causa incapacidad) y con enfermedades mentales como la demencia. Además concreta como población de riesgo de depresión y susceptible de cribado:

- Enfermedad de Parkinson, ya que se olvida frecuentemente su asociación con depresión.
- Demencia, dado que ambas enfermedades se parecen.
- Posparto (el cribado puede dar resultados positivos de hasta un 13%).

- Alcoholismo y drogas de abuso (en estos trastornos relacionados con el consumo de sustancias se complica decidir si la depresión es causa o consecuencia, pero, en cualquier caso, se puede tratar tanto la depresión como el trastorno relacionado con el consumo de sustancias).
- Víctimas de abuso.
- Enfermedad física como cáncer, enfermedad coronaria o diabetes.
- Dolor crónico.
- Estrés en el entorno familiar.
- Senectud.
- Aislamiento social.
- Síntomas inexplicados.

En resumen, la población diana de cribado va a solaparse en multitud de ocasiones con las personas con mayor probabilidad de sufrir depresión, especialmente aquellas personas que cumplen las características de mayor riesgo de padecer un episodio depresivo mayor según se ha visto en el Capítulo 1, a saber, ser mujer, estar soltero, divorciado o viudo, estar desempleado o de baja laboral, tener entre 30-50 años en hombres y entre 18-30 años en mujeres, o ser hiperfrecuentador de los servicios de atención primaria, entre otros.

2.3. Recomendaciones del uso del cribado de los trastornos depresivos

El uso del cribado en la detección de depresión ha sido recomendado por las guías clínicas pertenecientes a la *US Preventive Services Task Force* (2002), la *Canadian Task Force on Preventive Health Care* (2005) y la *UK National Institute of Clinical Excellence* o NICE (2006).

2.3.1. Recomendaciones de la *US Preventive Services Task Force* (USPSTF)

La evolución de las guías sobre cribado elaboradas por el Grupo de Trabajo de Servicios Preventivos de los Estados Unidos de América (*US Preventive Services Task Force* o USPSTF) ha sido dispar a lo largo de los años. En 1996 consideraba que los beneficios del uso del cribado de la depresión en AP no habían sido demostrados suficientemente como para recomendar su uso rutinario. En su última revisión, la guía de la USPSTF (2002) reconoce que, en base a la evidencia y la carga que supone esta enfermedad, hay que animar a los médicos de AP a que implementen el uso del cribado, el cual, “puede ser más productivo en pacientes con historia pasada de depresión, síntomas inexplicados somáticos, comorbilidad psicológica (como el trastorno de pánico o la ansiedad generalizada), abuso de sustancias o dolor crónico” p.122.

La guía además recomienda que un cribado positivo debe ser seguido por un diagnóstico preciso, un tratamiento efectivo y un seguimiento para

asegurar que los beneficios del cribado se lleven a cabo. Esta recomendación es valorada por la guía como grado B, es decir, recomienda que los clínicos de forma rutinaria proporcionen este servicio sobre el paciente. La USPSTF encuentra, al menos, datos empíricos razonables sobre que el uso del cribado aumenta los resultados positivos sobre la salud de los pacientes y concluye que los beneficios son superiores a los daños.

A su vez, la USPSTF también concluyó en las recomendaciones de 2002 que los datos empíricos para recomendar o contraindicar el cribado rutinario de depresión con cuestionarios estandarizados a niños o adolescentes era insuficiente.

La USPSTF encuentra pocos datos empíricos para recomendar una prueba de cribado sobre otra, de manera que se recomienda al clínico que escoja el método que más se adecue a sus preferencias personales, a la población a la que va dirigida y al servicio en el que se implementa.

Los daños potenciales del cribado incluyen la aparición de los falsos positivos, el inconveniente de desarrollar más diagnósticos, los efectos secundarios y los costes de medicar a pacientes incorrectamente identificados como depresivos, y el efecto adverso del etiquetado incorrecto. Según la USPSTF ninguno de los estudios revisados proporcionan datos empíricos que muestren estos efectos potenciales.

Por último, la guía de la USPSTF considera la idoneidad de utilizar programas integrados de cribado, los cuales incluirían dar *feedback* de los resultados, proporcionar psicoeducación, dar acceso al servicio de salud

mental, seguimiento telefónico y el compromiso institucional de aumentar la calidad del servicio.

2.3.2. Recomendaciones de la *Canadian Task Force on Preventive Health Care* (CTFPHC)

El Grupo de Trabajo Canadiense sobre Atención Sanitaria Preventiva (*Canadian Task Force on Preventive Health Care* o CTFPHC) encontró datos empíricos para excluir el cribado rutinario de depresión de las personas asintomáticas en su informe de 1994 (CTFPHC, 1994), pero sugería que los clínicos debían mantener un grado de alerta alto para diagnosticar esta patología entre sus pacientes.

Posteriormente, dicho grupo de trabajo revisó sus recomendaciones (MacMillan, Patterson y Whathen, 2005) concluyendo que en los servicios donde existe la posibilidad de proveer al paciente, además del cribado, de una retroalimentación y un tratamiento adecuados, los datos empíricos demuestran los resultados positivos del cribado rutinario y, por tanto, recomiendan su uso. En los demás casos no encuentran datos suficientes que recomienden su utilización rutinaria en AP. En la Tabla 2.1 se resumen los principales hallazgos y recomendaciones del Grupo de Trabajo Canadiense sobre Atención Sanitaria Preventiva en relación con el cribado de la depresión en AP.

Tabla 2.1 Resumen de las recomendaciones del Canadian Task Force on Preventive Health Care respecto al cribado de la depresión en atención primaria (MacMillan, Patterson y Whathen, 2005)

MANIOBRA	EFFECTIVIDAD	RECOMENDACIONES
Cribado en adultos de la población general en servicios con sistemas de <i>feedback</i> y tratamiento	Existe evidencia de que el cribado incrementa la precisión en la identificación de pacientes deprimidos. En estudios donde un sistema integral de cribado y seguimiento es posible, hay un aumento en los resultados.	Hay datos razonables para recomendar el cribado de la depresión en adultos en AP en los servicios en los que sea viable llevar a cabo programas integrados y derivar a los pacientes detectados a servicios de salud mental.
Cribado en adultos de la población general en servicios sin sistemas de <i>feedback</i> y tratamiento	Existe evidencia de que el cribado incrementa la precisión en la identificación de pacientes deprimidos. En estudios donde no hay un sistema integral de cribado y seguimiento, es menos posible un aumento en los resultados positivos.	No hay datos suficientes para recomendar en favor o en contra del uso del cribado en adultos en servicios de AP donde un seguimiento efectivo y un tratamiento no son posibles.
Cribado de depresión en niños y adolescentes de la población general	No se han encontrado estudios que examinen los resultados en niños o adolescentes de los cribados en atención primaria	No hay datos suficientes para recomendar en favor o en contra del cribado en adultos en servicios de AP donde un seguimiento efectivo y un tratamiento no son posibles.

2.3.3. Recomendaciones de la guía del *National Institute of Health and Clinical Excellence* (NICE)

El Instituto Nacional de Excelencia Clínica y de la Salud (*National Institute of Health and Clinical Excellence* o NICE) del Reino Unido en su guía clínica sobre depresión recomienda el uso de cribado de manera rutinaria en grupos de pacientes de alto riesgo (Goldberg, 2006).

En su revisión del 2009 (versión CG91, disponible en www.nice.org.uk/CG91) apuntan la importancia de estar alerta (especialmente en pacientes con historia de depresión o con enfermedad física crónica con discapacidad funcional asociada) y de preguntar dos cuestiones:

- En el último mes, ¿frecuentemente te has sentido preocupado por pensamientos como sentirte triste, deprimido o sin esperanza?
- En el último mes, ¿frecuentemente has sentido que tenías poco interés o placer en realizar cosas?

Si el paciente contesta positivo, entonces el médico de AP seguirá preguntando:

- En el último mes, ¿frecuentemente te has sentido inútil?
- En el último mes, ¿frecuentemente te has preocupado por tener una concentración pobre?
- En el último mes, ¿frecuentemente te has preocupado por tener pensamientos de muerte?

También recomiendan analizar el estado mental del paciente, su nivel de funcionamiento y sus dificultades sociales e interpersonales, así como el uso

de termómetro del estado de ánimo y/o la petición de información a familiares allegados en personas con dificultades lingüísticas, cognitivas o sensoriales. En los casos en los que exista sospecha de depresión, el NICE recomienda el uso de una medida validada para evaluar la depresión.

2.3.4. Recomendaciones de otras organizaciones o instituciones científico-profesionales

El Colegio Americano de Obstetras y Ginecólogos recomienda a clínicos de su área de especialización que estén atentos para detectar los síntomas de depresión y preguntar a los pacientes sobre posibles estresores psicosociales y sobre antecedentes familiares de depresión. La Academia Americana de Pediatría recomienda que los pediatras indaguen sobre la depresión en su rutina diaria de atención al adolescente, y la división 18 de la Asociación Médica Americana recomienda el cribado de depresión entre los adolescentes con riesgo, por ejemplo, con problemas familiares o uso de drogas o alcohol.

Por el contrario, la guía clínica del 2003 del *Institute for Clinical Systems Improvement* (ICSI) considera que no hay suficientes datos empíricos directos para recomendar la realización de cribados de depresión en la población general.

Con respecto a España, el Programa de actividades preventivas y de promoción de la salud (PAPPS) de la Sociedad Española de Medicina Familiar y Comunitaria (semFYC, 2001) no recomienda el cribado rutinario en población general, pero sí en población o situaciones de riesgo, a saber, en las

transiciones psicosociales y si existen síntomas o conductas que sugieran patología depresiva.

Ese grupo de PAPPS de la semFYC, en su revisión del 2009, consideran que la AP “puede desempeñar un papel relevante en el nuevo concepto de abordaje de los problemas de salud mental considerado por la OMS, en el que la prevención y el abordaje precoz son retos necesarios en los que el primer nivel de atención puede tener un papel clave; del mismo modo, la corresponsabilización de los profesionales de la salud mental y de atención primaria en el seguimiento de estos pacientes permitirán su tratamiento en contextos más integrados en la familia y la comunidad”. (Fernández-Alonso et al., 2009 p.22)

El grupo de PAPPS de la semFYC incluye en sus programas el “Diagnóstico precoz de la depresión y de los trastornos por ansiedad” tanto a nivel infanto-juvenil como en la edad adulta (semFYC, 2009). Para ello las recomendaciones que realiza son:

- Una detección precoz y adecuada de los trastornos psicopatológicos, seguido de un adecuado tratamiento psicosocial, ya que puede mejorar la calidad de vida y el pronóstico de los pacientes, además de disminuir la iatrogenia y reducir el gasto sanitario.
- La entrevista clínica constituye sin duda el mejor método diagnóstico y terapéutico para la detección de los trastornos psicopatológicos, si bien requiere una serie de habilidades por parte del entrevistador.
- La integración de algunos cuestionarios en el seno de la entrevista puede facilitar y mejorar la detección de psicopatología. Su propuesta en

ese sentido es la utilización de la Escala de Ansiedad y Depresión de Goldberg (EADG). En esta escala sólo se deben puntuar los síntomas que duren más de 2 semanas.

Es importante resaltar que la EADG no proporciona diagnóstico de «caso» psiquiátrico, sino tan sólo la probabilidad de serlo. En los pacientes en los que la EADG resulte positiva se recomienda realizar una entrevista diagnóstica semiabierta o dirigida más exhaustiva, teniendo en cuenta el componente biológico, la problemática y el soporte social y los aspectos emocionales. El diagnóstico debería confirmarse ajustándose a los criterios diagnósticos del CIE-10 o DSM-IV.

Como puede observarse recomiendan el uso de un cribado concreto para favorecer la detección precoz de la depresión en AP.

2.4 Controversias sobre el uso del cribado: resultados metaanalíticos

Como se ha visto en los epígrafes anteriores, existen diferentes puntos de vista sobre el uso de métodos de cribado de forma rutinaria en AP, y el tema se considera un tema polémico en los últimos años, por lo que es importante, más allá de consultar las guías clínicas elaboradas por distintas instituciones y organizaciones científico-profesionales, analizar los resultados de los estudios empíricos al respecto y, en concreto, los resultados de las revisiones o metaanálisis que se basan en dichos resultados.

En un estudio que evalúa sistemáticamente cuatro revisiones metaanalíticas sobre el uso del cribado de la depresión en AP (Hickie, Davenport y Ricci, 2002), se encuentra que dos de esas revisiones (Mulrow,

Williams y Gerely, 1995; Williams, Hitchcock, Cordes, Ramirez y Pignone, 2002) concluyen que existen varios instrumentos aceptables desde el punto de vista psicométrico, así como desde el punto de vista de su implementación práctica, para realizar cribados de la depresión en AP. Las otras dos revisiones se fijan básicamente en los resultados que tiene el realizar cribados de depresión en AP y ofrecen recomendaciones diferentes. Por una parte estarían las que parten de la *United States Preventive Services Task Force* (USPSTF, 2002), en las que se aboga por el uso del despistaje cuando es utilizado con la apropiada retroalimentación de la información sobre el diagnóstico por parte del facultativo, así como con el seguimiento adecuado con tratamientos basados en la evidencia. De otro lado, Gilbody, House y Sheldon (2001, 2008) se posicionan en contra del cribado, argumentando la falta de resultados positivos cuando el cribado es administrado de manera rutinaria.

Pygnone et al. (2002) realizan un metaanálisis de 7 estudios independientes con el objetivo de evaluar el efecto del cribado rutinario de depresión en la población adulta y concluyen que el riesgo relativo de seguir deprimido después de la intervención es del 0,87 (95% IC = 0,79-0,95). Dicho de otro modo, este resultado indica que el cribado de la depresión resulta en una reducción del 13% del riesgo. En puntuaciones absolutas, encontraron una reducción de un 9% en la proporción de pacientes con depresión crónica. Aun cuando estas reducciones pueden parecer pequeñas, el mismo trabajo analiza los datos de coste-efectividad y concluyen que el cribado de depresión tiene el potencial de producir ratios de coste-efectividad similares a los del cribado del cáncer de mama o de la hipertensión. Sin embargo, estos mismos autores reconocen la necesidad de investigar más para determinar la aportación del

cribado en estos resultados y para descubrir cuáles son los componentes de estos programas de prevención que son más efectivos. Otro aspecto a resaltar es la gran heterogeneidad de los estudios incluidos en el metaanálisis. Para mejorar en homogeneidad y así tener datos más precisos, realizan un análisis alternativo eliminando un estudio que había mostrado resultados muy positivos del cribado. Los datos encontrados son similares al estudio inicial, obteniendo un riesgo relativo para seguir deprimido tras el tratamiento de 0,90 (95% IC = 0,82-0,98) y una reducción de un 7% en la proporción de pacientes con depresión crónica, es decir, decrecen los resultados positivos ligeramente respecto al anterior análisis.

Algunos estudios como el de Dowrick, Bellon y Gómez (2000) sugieren que un retraso de la información obtenida por el cribado conducirá a una reducción de la efectividad.

En cuanto a las posiciones contrarias al uso del cribado de depresión de forma rutinaria, Gilbody et al. (2001, 2008) se plantean en sus estudios determinar el coste-efectividad clínica de los instrumentos de despistaje. Para ello consideran tres cuestiones:

1) ¿Mejora el reconocimiento de la depresión?

El estudio concluye que los cribados tienen un impacto límite sobre el reconocimiento general de la depresión (riesgo relativo = 1,38; intervalo de confianza del 95% = 1,04 a 1,83), aunque encontraron una heterogeneidad sustancial en esta medida de resultado. También hallaron que el cribado y la retroalimentación no tenían impacto alguno, independientemente de la puntuación inicial de la depresión, sobre la detección de la misma (riesgo

relativo = 1,00; intervalo de confianza del 95% = 0,89 a 1,13). En contraste, en los estudios bifase, donde primero se realiza un cribado y en el caso de que puntúe positivo, se introduce en el ensayo, los datos encontrados sugieren que este enfoque puede ser efectivo (riesgo relativo = 2,66; intervalo de confianza del 95% = 1,78-3,96).

2) ¿Mejora el tratamiento de la depresión?

Se encuentra una tendencia general a mostrar una tasa de intervención limítrofe superior entre los que recibieron retroalimentación de los instrumentos de despistaje (riesgo relativo = 1,35; intervalo de confianza del 95% = 0,98 a 1,85), aunque existía una heterogeneidad sustancial entre los estudios y los autores sugieren que este resultado dependió de un estudio sumamente positivo.

3) ¿Mejora el resultado de la depresión?

Gilbody et al. (2001, 2008) refieren que apenas encontraron estudios sobre el impacto de los instrumentos de detección de cribado sobre el resultado real de la depresión y que ninguna combinación estadística fue posible para evaluar este punto. Por otro lado, ningún estudio analizó la relación coste-efectividad del cribado.

En resumen, las conclusiones que aportan los metaanálisis de Gilbody et al. (2001, 2008) son que el uso del despistaje es un ejercicio costoso y que aparecen bajos beneficios en el incremento de la detección cuando éstos son usados en servicios no psiquiátricos.

Por su parte, el grupo de investigación australiano que, como se dijo anteriormente, realizó una evaluación sistemática de los metaanálisis previos

(Hickie et al., 2002), llegó a las siguientes conclusiones tras analizar los datos de las cuatro revisiones metaanalíticas mencionadas:

- Los instrumentos breves de cribado de la depresión tienen buenas propiedades psicométricas y son prácticos para su uso en servicios médicos generales.
- El cribado incrementa el reconocimiento e identificación de la depresión y, cuando es integrado en un compromiso de proveer coordinación y seguimiento tanto del diagnóstico como del tratamiento, los resultados clínicos se ven aumentados.

En la Tabla 2.2 se resumen las características y resultados principales de los cuatro estudios metaanalíticos más relevantes sobre cribado de la depresión en AP analizados por Hickie et al. (2002) y mencionados en los párrafos anteriores.

Tabla 2.2. Resumen de las características de las principales revisiones sobre el cribado de la depresión

Características	Estudios de revisión				
	Mulrow et al. (1995)	Gilbody et al. (2008)	Pignone et al. (2002)	Williams et al. (2002)	
Años revisados	1997-1994	1966-2004	1976-2001	1977-1999	
Factores examinados	Depresión; instrumentos de cribado; epidemiología; AP	Instrumentos de cribado de la depresión y retroalimentación (<i>feedback</i>) a los médicos	Cribado; proporción de reconocimientos y de tratamiento; resultados clínicos; AP	Cuestionarios de detección de depresión; proporción de reconocimiento; AP	
Número y tipos de estudio	9 estudios prospectivos en AP	12 estudios controlados aleatorios sobre la administración de cribados de depresión en AP u hospital general (<i>feedback</i> a los clínicos frente a no <i>feedback</i>). Se excluyen los estudios que utilizan cribado más otra estrategia de intervención.	7 estudios pilotos en AP (cribado rutinario frente a servicio habitual)	11 estudios prospectivos en AP	

Tabla 2.2. Resumen de las características de las principales revisiones sobre el cribado de la depresión (Continuación)

Características	Estudios de revisión			
	Mulrow et al. (1995)	Gilbody et al. (2008)	Pignone et al. (2002)	Williams et al. (2002)
Resultados	No se encontraron diferencias significativas entre instrumentos: sensibilidad general = 84%; especificidad = 72%	El cribado para la depresión administrado como rutina tiene un impacto mínimo sobre la detección de la depresión.	Sugieren que el cribado y <i>feedback</i> proveen de una reducción del 13% en el riesgo relativo de la depresión	No existen diferencias significativas entre los cuestionarios.
Conclusión	Bastantes cuestionarios son válidos con razonables características psicométricas para la identificación de pacientes con depresión. El cribado puede aumentar el reconocimiento de la depresión en AP.	No se deben seguir las recomendaciones de las guías de práctica clínica de adoptar el cribado por sí solo para mejorar la calidad de la atención sanitaria.	El cribado de depresión puede aumentar los resultados de los pacientes en AP, particularmente cuando el cribado está integrado en el propósito de mejorar el tratamiento y los sistemas clínicos (seguimientos coordinados en diagnóstico y tratamiento). El cribado de depresión con retroalimentación a los clínicos generalmente incrementa el reconocimiento de la depresión.	Encuentran bastantes cuestionarios con razonable rendimiento que son válidos para el uso por parte del clínico en la identificación y diagnóstico de pacientes con depresión.

En los últimos años se ha abogado por modelos de intensificación de la atención para la depresión, creando paquetes de intervención, tales como la “atención colaborativa”, que en general son efectivos (Gilbody et al., 2008). Prácticamente todos estos paquetes incluyen el cribado (Gray, 2004). Sería relevante analizar en qué grado el cribado aislado favorece este beneficio sobre la identificación del paciente depresivo.

Estos modelos multicomponentes que han mostrado buena efectividad estarían compuestos por:

1. Atención multiprofesional incluyendo el médico de AP, especialistas de salud mental y un gestor de caso (podría ser un profesional o paraprofesional que tuviera contacto rutinario con el paciente con el fin de apoyar a nivel psicosocial o con el seguimiento del tratamiento).
2. Una planificación en el tratamiento que abarcara tanto terapia psicológica breve como tratamiento farmacológico cuando fuera apropiado.
3. Seguimiento programado con recogida sistemática de los datos para realizar supervisión y favorecer la toma de decisiones sobre planes de tratamiento.

En resumen, se trata de proveer un modelo de cuidados coherente, integrado e individualizado para la persona con patología depresiva (Boarman y Walters, 2009).

De hecho, los últimos estudios (Magruder y Yeager, 2010) refieren el cambio tan importante en relación al cribado en depresión en AP en los últimos 20 años, tornándose el médico de atención primaria como una figura clave en

la detección y tratamiento de la depresión. Las miras al futuro irían en la línea de conseguir instrumentos cada vez más eficientes (más cortos con mejores propiedades psicométricas) e ir introduciendo tecnologías (Ej. que se puedan realizar en la sala de espera con monitores que recojan de manera confidencial los datos de cada paciente y que incluso puedan realizarse de manera sistemática cada cierto tiempo para la monitorización de la respuesta al tratamiento).

Capítulo 3:

Instrumentos de cribado de los trastornos depresivos en atención primaria

“¡Triste época la nuestra!

Es más fácil desintegrar un átomo que un prejuicio”

A. Einstein

3.1. Instrumentos de cribado de los trastornos depresivos en atención primaria en España

En España no se utiliza de forma rutinaria en AP un instrumento estandarizado de despistaje para la detección de los trastornos depresivos. Como aparecía en la cita de la portada de Einstein, revocar la percepción derivada de un prejuicio, o incluso de la denominada “impresión u ojo clínico” no es sencillo. En capítulos anteriores se ha visto las dificultades asociadas a un diagnóstico preciso cuando se usa casi de manera exclusiva la entrevista clínica para realizar el diagnóstico de depresión. La determinación de un diagnóstico clínico debería seguirse de la entrevista clínica, el conocimiento del facultativo y ser complementada por pruebas objetivas y empíricamente validadas. Aunque varios instrumentos de cribado han demostrado aceptables propiedades psicométricas en población española, como el BDI (Sanz, Navarro y Vázquez, 2003), el HDRS (Bobes et al., 2003), el SDS (Conde, 1976; Conde y Franch, 1984) o el CES-D (Soler, 1997), éstos no han sido concebidos específicamente para la detección específica de la depresión en AP. Otros, como el PRIME-MD y su versión abreviada el PHQ (Baca et al. 1999; Diez-Quevedo, Rangil, Sánchez-Planell, Kroenke y Spitzer, 2001) o la HAD (Caro e Ibáñez, 1992; Herrero et al., 2003; Quintana et al., 2003), sí que fueron diseñados para su utilización en AP, obteniendo resultados aceptables en Estados Unidos de América, pero sus propiedades psicométricas no han sido examinadas en profundidad en España (Reuland, Cherrington y Watkins, 2009).

Un aspecto fundamental a la hora de evaluar la depresión, como cualquier otro constructos psicopatológico, es valorar las características de

calidad y utilidad del instrumento con el que se va a realizar dicha evaluación. La proliferación de cuestionarios y escalas desde los años 80 del siglo pasado ha sido ingente, por lo que es necesaria la búsqueda de criterios que faciliten la mejor selección del instrumento adecuado en función de las necesidades de la evaluación. De manera breve, siguiendo a Marquis (1998) y con el fin de facilitar el análisis de los cuestionarios propuestos vamos a detallar algunas directrices para la descripción e interpretación de un test.

- Resumen de los contenidos del cuestionario:

Usando un formato estandarizado que podría incluir aspectos como: nombre y autor, objetivos, población diana, conceptos medidos (descripción de dimensiones), características de aplicación (ítems, periodo de retest, tipo de escalas usadas, tiempo de ejecución), validez, formas de administración, idiomas, disponibilidad.

- Informe del desarrollo y validación del cuestionario:

- Modelos conceptuales y de medida.

- Fiabilidad.

- Validez.

- Sensibilidad.

- Datos específicos de interpretación:

- Información sobre contenidos y constructos.

- Datos estadísticos utilizados para evaluar diferencias y cambios.

- Baremos

- Niveles de referencia obtenidos a través de niveles de gravedad reconocidos y de adecuados tamaños muestrales.

- Criterios dependientes, tales como la puntuación global de condiciones y cambios.

- Criterios independientes, tales como los recursos utilizados en diferentes rangos de puntuación.

- Criterios predictivos incluyendo valores de diagnóstico y pronóstico.

- o Informe de resultados:

- Descripción de la muestra.

- Línea base y distribución de puntuaciones de seguimiento incluyendo los valores perdidos.

- Distribución de cambios en las puntuaciones.

- Tipo de análisis y test usados.

- Inclusión de valores perdidos.

- Valores p si son relevantes.

- Interpretación de los datos.

La consideración de estos indicadores permite asegurar la idoneidad del instrumento con que se va a realizar la medida de la depresión. La tendencia actual es hacer énfasis en la evaluación de los test que se han editado previamente o realizar las directrices para los que están en curso. De hecho, desde el 2000, el Colegio Oficial de Psicólogos de España, siguiendo las directrices de la Comisión Europea de Test (EFPA-SCTT) adaptan dicho modelo al contexto español (Prieto y Muñiz, 2000). En 2010 inician el proceso

de evaluación de 10 test ampliamente utilizados por los psicólogos así como por el interés mostrado por los editores para someter sus test a dicha evaluación. La idea es ir ampliando el proyecto a la evaluación de la totalidad de las pruebas (Muñiz, Fernández-Hermida, Fonseca-Pedrero, Campillo-Álvarez y Peña-Suárez, 2011). En este sentido, en los próximos epígrafes se describen los cuestionarios de mayor relevancia para el cribado de la depresión en España tanto a nivel clínico como de investigación, utilizándose una versión resumida de los indicadores anteriores.

3.1.1.- Escala de Depresión de Hamilton (HDRS)

Autor: M. Hamilton

Referencia: Hamilton, M. (1969). A rating scale for depression. *Journal of Neurosurgery and Psychiatry*, 23, 56-62.

Contenido que evalúa: La existencia y gravedad de síntomas depresivos.

Nº de ítems: Existen más de 20 versiones, siendo las más populares las de 6, 7, 17, 21 y 25 ítems.

Administración: Heteroaplicada

Descripción:

- La HDRS evalúa el perfil sintomatológico y la gravedad del cuadro depresivo.
- Ha de ser administrada por un clínico entrenado.
- El marco de referencia temporal es en el momento de la entrevista, excepto en algunos ítems, como los de sueño en los que se exploran los dos días previos.

Aplicación: A partir de entrevista individual en un tiempo aproximado de 30 minutos.

Edad: Adultos

Corrección e interpretación:

- Proporciona una puntuación global de gravedad del cuadro y una puntuación en tres factores o índices: melancolía, ansiedad y sueño.
- La puntuación global se obtiene sumando las puntuaciones de cada ítem.
- Existen distintos puntos de corte o normas de interpretación. Inicialmente se crea con 21 ítems pero es el propio Hamilton quien considera que los cuatros últimos ítems no se deben computar por la baja frecuencia de aparición en los trastornos depresivos (variación diurna, despersonalización/desrealización, síntomas paranoides y síntomas obsesivos/compulsivos).
- Los autores del instrumento proponen para la interpretación de la puntuación global los puntos de corte que se recogen en la Tabla 3.1.

Tabla 3.1 Puntos corte del instrumento de depresión Hamilton

<i>Punto de corte</i>	<i>Gravedad</i>
0–6	Depresión ausente
7–17	Depresión leve
18–24	Depresión moderada
A partir de 25	Depresión grave

Furukawa (2010) plantea unos puntos de corte alternativos que se presentan en la Tabla 3.2.

Tabla 3.2. Puntos de corte del instrumento de depresión de Hamilton según Fukurawa (2010)

<i>Punto de corte</i>	<i>Gravedad</i>
0–3	Depresión ausente
4–7	Límite con patología depresiva
8–15	Depresión leve
16–26	Depresión moderada
A partir de 27	Depresión grave

- El criterio de inclusión más utilizado en los ensayos clínico es una puntuación ≥ 18 (depresión clínica) y ≥ 9 (sintomatología depresiva).
- Las puntuaciones en cada uno de los índices se obtiene sumando las puntuaciones de los ítems que lo constituyen: melancolía (ítems 1, 2, 7, 8, 10, 13); ansiedad (ítems 9-11) y sueño (ítems 4-6). No existen puntos de corte definidos para la puntuaciones en estos índices.

Propiedades psicométricas:

- Fiabilidad: En la Tabla 3.3 se presenta un resumen de los índices psicométricos de fiabilidad de las diferentes versiones de la HDRS en su adaptación española.

Tabla 3.3. Índices de fiabilidad HDRS (Bobes et al., 2003)

<i>Propiedades</i>	<i>HDRS-7</i>	<i>HDRS-17</i>	<i>HDRS-21</i>
Consistencia interna α de Cronbach	0,59	0,74	0,78
Fiabilidad test-retest	0,93	0,93	0,92
Fiabilidad entre observadores	0,97	0,99	0,99

- En un metaanálisis realizado a través de la generalización de la fiabilidad (medida con el α de Cronbach), el HDRS alcanza unos índices de fiabilidad que varían de $\alpha = 0,41$ a $0,89$ (López-Pina, Sánchez-Meca y Rosa-Alcázar, 2009). Realizando la media corregida (teniendo en cuenta la importante variabilidad entre las diferentes versiones, métodos de estudio, poblaciones...) se encuentra un valor medio de $0,80$ ($95\% \text{ IC} = 0,79-0,81$). En este sentido, se obtienen datos similares comparándolos con estudios anteriores realizados a través de la misma técnica en relación con el BDI ($0,84$; Yin y Fan, 2000) y con el CES-D ($0,88$; O'Rourke, 2004).
- Morriss, Chatwin y Baldwin (2003) analizan la concordancia interjueces de la versión de 17 ítems en el contexto de AP. El coeficiente intraclase y el coeficiente de concordancia para la puntuación total fue en ambas del $0,95$, con un 95% intervalo de referencia (para la diferencia entre pares) de jueces que iba de $-3,31$ a $3,69$. La fiabilidad interjueces no varía con la gravedad de la depresión o por sesgos de los diferentes jueces. El coeficiente kappa para los ítems individuales fue bueno, aproximadamente de $0,6$. Los autores refieren su buena capacidad tanto para el apoyo al

diagnóstico como para la inclusión de pacientes en ensayos clínicos (Morriss, Chatwin y Baldwin, 2003).

- Validez de criterio: Respecto al área bajo la curva ROC, al compararse en las 3 versiones se encontró que la HDRS-6 discriminaba mejor entre pacientes depresivos y no depresivos que la HDRS-17 y la HDRS-21 (Bobes et al., 2003). Dado que la HDRS-21 no se diferencia de la HDRS -17 en su capacidad de discriminación, al igual que la HDRS-7, y que ésta no representa un modelo significativamente distinto de las escalas de 17 y 21 ítems, la elección de una u otra puede hacerse según el fin que se quiera utilizar.
- Validez convergente HDRS: 0,67-0,76 (en relación con el BDI o la SDS de Zung)
- Validez discriminante HDRS: las tres versiones del HDRS son capaces de discriminar de manera significativa entre pacientes con y sin trastorno depresivo. (Bobes et al., 2003)
- Sensibilidad al cambio: elevada (medida por el tamaño del efecto; d de Cohen = 1,6).

Furukawa (2010) plantea la siguiente guía para valorar el cambio clínicamente significativo:

- 18 puntos ó 73%: Mejora muy relevante
- 11–17 puntos ó 46–72%: Mejora relevante
- 5–10 puntos ó 15–45%: Mejora mínima
- 0–4 puntos ó 5–14%: Sin cambios

- Estructura factorial: Diversos estudios encuentran soluciones factoriales que van desde 2-8 factores. Parece que en todos se repiten tres factores (depresión, dificultades del sueño y ansiedad/agitación). Gibbons, Clark y Kupfer. (2002) encuentra en su estudio una solución de cinco dimensiones (ánimo deprimido, culpa, trabajo e intereses, agitación, ansiedad ideica, ansiedad somática y pérdida de libido).

Ventajas e inconvenientes:

- Existen varios guiones de entrevista para la realización del HDRS pero ninguno oficial. En inglés la versión estructurada más actual es la GRID-HAMD creada por Williams. et al.(2008)
- El HDRS ha sido uno de los métodos más populares durante 50 años en la práctica clínica de la salud mental. (López-Pina, Sánchez-Meca y Rosa-Alcázar, 2009) y uno de los instrumentos más utilizados para medir la gravedad en depresión (Kriston y Wolff, 2011).
- A lo largo de estos años, las críticas han ido en aumento. En un estudio (Bagby, Ryder, Schuller y Marshall, 2004) se analiza las propiedades psicométricas del HDRS en las publicaciones en MEDLINE para el período de 1979-2003 y se concluye que:
 - o La fiabilidad medida a través de su consistencia interna es adecuada, pero varios ítems contribuyen de manera pobre a las medidas de gravedad de la depresión; en otros estudios aparece baja fiabilidad test-retest.
 - o La validez de contenido es baja, no cubre varios de los criterios del DSM-IV (APA, 1994) como, por ejemplo, dificultades de concentración y

sentimiento de incapacidad. Es especialmente reseñable la no existencia de un ítem para la anhedonia, uno de los dos criterios necesarios para el diagnóstico del episodio depresivo mayor.

- La validez convergente y discriminante es adecuada.
- La estructura factorial de la HDRS es multidimensional, pero con pobres replicaciones en diferentes muestras.
- Parece que la extensión de la HDRS-21 y la HDRS-17 así como el tiempo elevado que se emplea en aplicarla ha sido una de sus principales limitaciones en su uso rutinario en contextos médicos. Bobes et al. (2003) abogan por el uso de su versión corta de 6 ítems como un instrumento muy adecuado para el médico de primaria o personal no especialista en salud mental.
- De hecho, pese a ser uno de los instrumentos más utilizados a nivel de investigación clínica, la variabilidad en criterios sobre la gravedad en relación al HDRS varía de manera sustancial en diferentes instituciones, lo que hace que sea difícil la comparación y generalización de resultados en las investigaciones. En este sentido, es esclarecedor el cuadro gráfico que realizan Kriston y Wolf (2011) para describir esta situación y que se recoge en la Tabla 3.4.

Tabla 3.4. Niveles de depresión de la HDRS recomendados por diferentes autores

Puntuaciones de la Escala de Hamilton																																	
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 ≥31																																	
Bech,(1996)	No deprimido					Ligera					Menor que mayor					Mayor										Grave							
	No deprimido																																
APA, (2000)	No deprimido					Ligera					Moderada					Grave					Muy grave												
Furukawa, (2007)	Asintomático					Ligera					Moderada										Grave												
NICE (2009)	No deprimido					Subclínica					Ligera					Moderado					Grave												
Baer (2010)	No deprimido					Ligero										Moderado										Grave							

- Pero no solo la variabilidad es a nivel de gravedad, sino que existen múltiples adaptaciones, modificaciones y traducciones. De tal forma que un estudio que recabó las diversas formas originales usadas en diferentes investigaciones publicadas en revistas de impacto concluyó que apenas coincidían, es decir se estaban utilizando diferentes versiones, sin tener en cuenta las consecuencias a nivel de sus respectivas propiedades psicométricas (Fukorawa, 2010).
- En conclusión, parece que actualmente habría que revisarlo y actualizarlo, al menos para que se siguiera considerando como un instrumento de referencia. La importante variabilidad en las concepciones y categorizaciones del HDRS puede tener consecuencias importantes como la consideración de determinadas actuaciones y recomendaciones en función de la institución ante una misma gravedad en la patología. Es una escala poco usada en población no clínica. Históricamente es una escala

que se ha utilizado en ámbito médico y de ensayos clínicos, lo cual es llamativo dado que aparece un elevado solapamiento con síntomas de ansiedad y con síntomas somáticos, lo que dificultaría su administración en enfermos físicos.

3.1.2.- Escala de Depresión del Centro de Estudios Epidemiológicos (CES-D)

Autor: L. S. Radloff

Referencia: Radloff, L. S. (1977). The CES-D scale: a self-report depression scale for research in general population. *Applied Psychological Measurement*, 1(3), 385-401.

Evalúa: Presencia y gravedad de síntomas depresivos.

Nº de ítems: 20 ítems.

Administración: Autoaplicada

Descripción:

- La escala CES-D contiene 20 ítems que evalúan la frecuencia de los síntomas (0-3).
- El marco temporal de evaluación de los síntomas es de la semana anterior.
- El tiempo de aplicación de es 5-10 minutos.
- Versión española de Soler, Pérez Sola, Puigdemot, Pérez-Blanco y Figueres (1997).
- El punto de corte usualmente utilizado es ≥ 16 (Magruder y Yeager, 2010).

Propiedades psicométricas:

- Fiabilidad: $\alpha = 0,85-0,90$
- Método de las dos mitades: 0,89
- Validez convergente: correlación CES-D y BDI = 0,79, y correlación CES-D y HDRS = 0,69.
- Sensibilidad: 81%; 93% (85-97)
- Especificidad: 72%; 69% (65-74)
- Curva ROC (95%CI): 0,89 (0,85-0,92) (Citado en Magruder y Yeager, 2010)
- Validez de criterio: Alta concordancia interevaluadores $kappa = 0,95$.

Más recientemente se creó una versión reducida de 7 ítems que parece una forma de cribado aceptable para el uso en AP. La versión reducida del CESD (CESD-7) se crea a partir de la versión reducida del CESD a nueve ítems de Santor y Coyne (1997; citado en Herrero y Gracia, 2007). Estos autores analizan los ítems que resultan más discriminativos para la depresión y aparece una solución de 7 ítems. Se incluyen los ítems más relacionados con los síntomas comunes de la depresión: ánimo disfórico (3, 6 y 18), motivación (7), concentración (5), pérdida del placer (16) y dificultades con el sueño (11).

Los índices psicométricos del CESD-7 son aceptables, presentando una satisfactoria fiabilidad medida a través de su consistencia interna de 0,82 y fiabilidad test-retest (los resultados se mantienen estables a los seis meses) así como una solución unifactorial y una adecuada validez de constructo (Herrero y Gracia, 2007). No especifican puntos de corte para facilitar la discriminación entre personas depresivas y no depresivas.

Ventajas:

-Fácil de aplicar, sensible y específica en *screening* en población psiquiátrica y estudios epidemiológicos en población general.

Desventajas:

-Escasos estudios analizando las propiedades psicométricas de la adaptación española del CES_D.

-Baja especificidad de la prueba.

3.1.3.- Test de dos preguntas

Autores: B. Arroll, N. Khin y N. Kerse

Referencia: Arroll, B., Khin, N., y Kerse, N. (2003). Screening for depression in primary care with two verbally asked questions: cross sectional study. *BMJ*, 327, 1144–1146.

Evalúa: Detección de depresión.

Nº de ítems: 2 ítems.

Administración: Heteroaplicada verbalmente.

Propiedades psicométricas:

- Validez de criterio: Los autores refieren que el test de dos preguntas es el cribado más corto existente con aceptables propiedades psicométricas. En su estudio comparan con una versión computerizada de la CIDI (*Composite International Diagnostic Interview*; WHO, 1990) obteniendo una sensibilidad 97% (95% IC = 83%-99%) y una especificidad del 67% (95% IC = 62%-

72%). Al analizar los índices de valor predictivo positivo comparados con la CIDI, se encuentra un elevado índice de falsos positivos, lo cual compromete de manera relevante la calidad de la prueba. Henckel, Megl, Coyne y Kohnen (2004) consideran que cuando se utilizan dos preguntas los resultados son más favorables, respecto al cribado de una pregunta, apareciendo propiedades aceptables respecto a test más largos

Ventajas e inconvenientes:

En los últimos años aparecen autores a favor del uso de cribados ultracortos. Mitchell y Coyne (2007) en su metaanálisis consideran que los cribados ultracortos (entre uno y cuatro ítems) tienen una capacidad significativa (medida a través del tamaño del efecto $d = 5,46$) para identificar deprimidos/no deprimidos comparándolo con el azar. En esta línea, el NICE (*National Institute for Health and Clinical Excellence*) recomienda el uso de cribados de una y dos preguntas para la depresión sin embargo concreta que estos métodos pueden ser usados como despistaje inicial, pero una vez que han puntuado positivo, deben ser seguidos por un método de cribado más preciso. El problema principal que plantean este tipo de cribado es la baja especificidad de la prueba, por ejemplo, Mitchell y Coyne (2007) refieren que el test de una pregunta sólo consigue identificar a tres de cada diez casos con depresión en AP, lo cual supone unos índices de discriminación inaceptables.

- Parece que son necesarios más estudios en la línea de mejorar la especificidad de la prueba dada la elevada tasa de falsos positivos que aparecen en este tipo de instrumentos.

3.1.4. ESCALA AUTOAPLICADA PARA LA MEDIDA DE LA DEPRESIÓN DE ZUNG (SDS)

Autor: W. Zung

Referencia: Zung, W. (1965). A self-rating depression scale. *Archives of General Psychiatry*, 12, 63-70.

Evalúa: Existencia y gravedad de síntomas depresivos.

Nº de ítems: 20 ítems.

Administración: Autoaplicada

Descripción:

- Se trata de una escala breve, en la que el paciente debe indicar la frecuencia con la que experimenta cada uno de los síntomas explorados.
- Consta de 20 ítems que se agrupan en 4 factores:
 - Factor depresivo
 - Factor biológico
 - Factor psicológico
 - Factor psicosocial
- Cada ítem responde según una escala tipo Likert de 4 valores (rango de 1-4) que hace referencia a la secuencia de la presentación de los síntomas.
- El marco de referencia temporal es el momento actual. En la versión española es la última semana.
- Para evitar el sesgo de aquiescencia en las respuestas, se alternan ítems redactados en sentido positivo con ítems en sentido negativo.

- Tiempo de aplicación: 5 minutos aproximadamente.
- Esta escala está adaptada y validada al español (Conde, Escribá e Izquierdo, 1970; Conde y Franch, 1984).

Corrección e interpretación:

- Proporciona una puntuación total de gravedad que es la suma de las puntuaciones de los 20 ítems. Los ítems redactados en sentido negativo son: 1, 3, 4, 7-10, 13, 15 y 19. Los ítems redactados en sentido positivo son: 2, 5, 6, 11, 12, 14, 16-18 y 20.
- La puntuación oscila entre 20 y 80; a mayor puntuación, mayor gravedad del cuadro depresivo. Existe también un sistema de puntuación porcentual para la SDS basado en la transformación de las puntuaciones brutas en porcentajes sobre la puntuación bruta máxima de la escala.
- Se utilizan de manera prioritaria dos sistemas de puntos de corte que se recogen en las Tablas 3.5 y 3.6, uno basado en las puntuaciones brutas (Tabla 3.6) y otro basado en las puntuaciones porcentuales (Tabla 3.5).

Tabla 3.5. Puntos de corte de la SDS (Conde y Chamorro, 1973)

<i>Punto de corte</i>	<i>Gravedad de la depresión</i>
≥ 50%	Depresión subclínica y variantes normales
≥ 60%	Depresión moderada
≥ 70%	Depresión grave

Tabla 3.6. Puntos de corte de la SDS (Aragonés, Masdeu, Guasch y Borrás, 2001)

<i>Punto de corte</i>	<i>Gravedad de la depresión</i>
20–35	Depresión ausente
36–51	Depresión subclínica y variantes normales
52–67	Depresión moderada - grave
68–80	Depresión grave (opcionalmente hospitalizable)

Propiedades psicométricas:

- Fiabilidad: La consistencia interna [$\alpha = 0,78$; $r_{x,x'}(\text{corregida}) = 0,80$] es suficientemente alta, al igual que la fiabilidad test-retest ($r = 0,80$).
- Validez de contenido: Dudosa, no hay ítems de retardo, culpa o somáticos.
- Validez convergente: Se han encontrado correlaciones de la SDS con la HDRS iguales a 0,58, y correlaciones con la escala de depresión del MMPI (MMPI–D) que oscilan entre 0,59 y 0,70.
- Concordancia interobservadores: Kappa = 0,70
- Validez de criterio: Dudosa. Para un punto de corte $\geq 50\%$ aparece un valor predictivo positivo (VPP) de 0,35 y un valor predictivo negativo (VPN) de 0,99.

En el estudio de Aragonés, Masdeu, Cando y Coll (2001), se presentan buenos resultados para el uso de la escala SDS en AP. Los datos que obtuvieron de sensibilidad y especificidad para detectar ambos diagnósticos fueron de 0,95 (IC 95%: 0,87-1) y de 0,74 (IC 95%: 0,68-0,79). El área bajo la curva ROC fue de 0,93. Consideran que un punto de corte equilibrado puede

ser una puntuación en la SDS $\geq 54\%$, el cual, muestra una sensibilidad de 0,85 y una especificidad de 0,86. Estos resultados se aproximan a los obtenidos en otros estudios sobre la validez del SDS en pacientes de AP. Por ejemplo, Zung et al. (1967) hallaron, para detectar depresión mayor, una sensibilidad del 0,97 y una especificidad de 0,63 para el punto de corte de la SDS $\geq 50\%$ (citados en Aragonés et al., 2001).

Como ventajas aparece que es instrumento útil y rápido para la detección de la depresión. En la práctica clínica, considerando que se trata de una escala autoaplicada, se plantean problemas de comprensión y registro adecuado de las respuestas.

3.1.5.- Evaluación de los Trastornos Mentales en Atención Primaria (PRIME-MD Primary Care Evaluation of Mental Disorder)

Autor: R. L. Spitzer, J. B. Williams, K. Kroenke, deGruy, F. V. III, Hahn, S. R., et al.

Referencia:

- Spitzer, R. L., Williams, J. B., Kroenke, K., Linzer, M., deGruy, F. V. III, Hahn, S. R., et al. (1994). Utility of a new procedure for diagnosing mental disorders in primary care. *JAMA*, 272, 1749-1756.
- Adaptación española: Baca, E., Saínz, J, Agüera, L., Caballero, L., Fernández Liria, A., Ramos, J., Gil, A., Madrigal, M., y Porras, A. (1999). Validación de la versión española del PRIME-MD: Un procedimiento para el

diagnóstico de los trastornos mentales en Atención Primaria. *Actas Españolas de Psiquiatría*, 27(6), 375-383.

Evalúa: Presencia de síntomas de trastornos del estado de ánimo, trastornos de ansiedad, trastornos de la conducta alimentaria y alcoholismo.

Nº de ítems: 25 ítems para el cuestionario (CP).

Administración: Autoaplicada (fase del paciente: CP) y heteroaplicada (fase del clínico: GEM).

Descripción:

- El PRIME-MD (*Primary Care Evaluation of Mental Disorder*) se diseñó con el objetivo de diagnosticar trastornos específicos en AP usando criterios DSM-III-R y DSM-IV.
- Es un instrumento que combina un cuestionario y una entrevista con preguntas formuladas de forma dicotómica (SI/NO). Existe un último ítem que evalúa el estado de la salud percibida por el paciente.
- Se propone usarlo en dos fases:
 - o Fase 1ª. El paciente completa un cuestionario (CP) de 25 ítems, que actuaría de *screening* y que se compone de 5 módulos acerca de los siguientes trastornos: estado del ánimo, ansiedad, somatoforme, conducta alimentaria, abuso y dependencia de alcohol.
 - o Fase 2ª. Se compone de una guía de evaluación para el facultativo (GEM), la cual contiene información adicional de los mismos 5 módulos para profundizar en las áreas donde el paciente ha puntuado positivo en el CP.

- **Tiempo de aplicación:** 11 minutos

Propiedades psicométricas:

- Sensibilidad: 81%
- Especificidad: 66%
- Valor predictivo positivo:0,62
- Valor predictivo negativo: 0,83
- Concordancia: $\kappa=0,45$

En general, la especificidad es moderada, lo que facilita un aumento de los falsos positivos. Si se analiza por separado:

- o En los trastornos del estado de ánimo la sensibilidad y la especificidad son adecuadas; se considera útil el uso de éste módulo por separado en la detección de dichos trastornos.
- o Trastornos de ansiedad: buena sensibilidad en detrimento de la especificidad.
- o Trastornos somatomorfos: este módulo no ha mostrado validez. Especificidad y concordancia poco adecuadas.
- o Abuso y dependencia de alcohol: especificidad y sensibilidad adecuadas, pero modesta concordancia del CP respecto al GEM.
- o Trastornos de la Conducta Alimentaria: se necesitan mayores estudios en español para evaluar la validez en estos trastornos.

Existe una versión de dos ítems PHQ_2 (Magruder y Yeager, 2010)

- Sensibilidad: 96%

- Especificidad: 57%

Como puede observarse y en la línea de otros cribados ultracortos adolece de una especificidad baja.

Desventajas:

El tiempo de aplicación es alto. Se ha usado más en investigación que en la práctica clínica.

3.1.5.1.- Cuestionario de Salud del Paciente–9 (PHQ-9)

Autor: R. L. Spitzer, K. Kroenke y J. B. Williams

Evalúa: Presencia y gravedad de síntomas de trastornos del estado de ánimo.

Nº de ítems: 9 ítems + 1 ítem de funcionamiento.

Administración: Autoaplicada

Referencias:

- Spitzer, R. L., Kroenke, K., y Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *JAMA*, 282, 1737-1744.
- Kroenke, K., y Spitzer, R. L. (2002). The PHQ-9: A new depression and diagnostic severity measure. *Psychiatric Annals*, 32, 509-521.

Descripción:

- El Cuestionario de Salud del Paciente–9 (*Patient Health Questionnaire–9*) es el módulo de trastornos del estado de ánimo del PRIME-MD, creado a partir

de los criterios del trastorno depresivo mayor del DSM-IV. Se usa cada vez más como método de cribado en depresión.

- El tiempo de aplicación se reduce a aproximadamente 4 minutos.
- 4 categorías de respuesta que permite ver un cambio sintomático a lo largo del tiempo.

Corrección e interpretación:

- Su puntuación máxima es de 27.
- Se detecta depresión si puntúa en más de 5/9 ítems en la categoría “más de la mitad de los días” en las últimas dos semanas, siendo uno de los ítems sentimiento de tristeza o anhedonia.
- Otros tipos de depresión son diagnosticados si puntúa en 2, 3 ó 4 ítems, en las últimas dos semanas, siendo uno de los ítems sentimiento de tristeza o anhedonia. Tienen especial consideración con el ítem de suicidio: a no ser que sea elegido “nunca” como respuesta, puntuaría siempre, independientemente de la frecuencia y de la duración.

Tabla 3.7. Puntos de corte del PHQ-9 según Kroenke et al. (2001)

<i>Puntuación</i>	<i>Gravedad en depresión</i>
5	Depresión Ligera
10	Depresión Moderada
15	Depresión Moderada- grave
20	Depresión Grave

- Los puntos de corte para interpretar las puntuaciones de los pacientes se muestran en la Tabla 3.7. Los autores consideran como el punto de corte más recomendado una puntuación superior o igual a 10.
- Según los autores, previo a la realización de un diagnóstico, el clínico descartaría depresión debida a enfermedad física, duelo e historia de episodio maniaco. Por último se realizaría una pregunta relativa al nivel de funcionalidad del paciente.
- La diferencia clínicamente significativa se establece en 5, es decir se consideraría una respuesta al tratamiento relevante a aquellas personas que puntuasen 5 puntos menos en el postratamiento (Lowe, Unutzer, Callahan, Perkins y Kroenke, 2004).

Propiedades psicométricas:

- Fiabilidad: En cuanto a sus índices psicométricos, Spitzer et al. (1999) apuntan una fiabilidad excelente medida a través de su consistencia interna ($\alpha = 0,89$). La correlación test-retest a una semana también mostró resultados positivos con un valor de 0,92, lo que también es calificado como excelente (Pinto-Meza, Serrano-Blanco, Peñarrubia, Blanco y Haro, 2005).
- Validez de criterio: En su metaanálisis con 14 estudios, Gilbody et al. (2007) obtuvieron una sensibilidad de 0,80 y una especificidad de 0,92. Cabe resaltar la importante heterogeneidad de los estudios incluidos en el metaanálisis, ya que eran difícilmente equiparables en cuanto al punto de corte, dispositivo asistencial, tipo de muestra, etc. Siguiendo con su validez de criterio, Kroenke et al. (2001) consideran que el PHQ-9 diferencia adecuadamente entre depresión subclínica, patología y ausencia de

depresión. Además consideran que correlaciona linealmente con medidas como calidad de vida, días de incapacidad, visitas al clínico y autorregistro de dificultades debidas a los síntomas. Utilizando como criterio estándar una entrevista estructurada de trastornos mentales realizada por un profesional (basado en la SCID), una puntuación > 10 en el PHQ, obtiene una sensibilidad del 88% y una especificidad del 88% para el trastorno depresivo mayor.

- Validez convergente: Diez-Quevedo, Rangil, Sanchez-Planell, Kroenke y Spitzer. (2001) realizan un estudio sobre el PHQ sobre una muestra española y señalan que el análisis de la parte correspondiente al módulo de la depresión (PHQ-9) correlacionaba de manera adecuada con la entrevista estructurada realizada por el profesional clínico (0.71) y con la puntuación total del BDI (0.76).

En un estudio de 2005 validaron el procedimiento de administración vía telefónica obteniendo resultados similares a los de su aplicación autoadministrada (Pinto-Meza et al., 2005).

Ventajas y desventajas:

- El PHQ es eficiente, requiriendo un menor tiempo de aplicación y corrección respecto a su versión más larga (dado que se trata del modulo de depresión del PRIME-MD).
- Evalúa frecuencia y gravedad de los síntomas.
- Como punto fuerte está el importante valor predictivo del último ítem, que proporciona un potente indicador de la probabilidad de diagnóstico psiquiátrico o del funcionamiento general del paciente.

- Reuland et al. (2009), en un estudio donde comparan la precisión de distintos instrumentos de cribado como son el *Center for Epidemiologic Studies-Depression Scale* (CES-D), el *Patient Health Questionnaire-9* (PHQ-9), el *Beck Depression Inventory-II* (BDI-II), la *Geriatric Depression Scale* (GDS), la *Edinburgh Postnatal Depression Scale* (EPDS), y el *Postpartum Depression Screening Scale* (PPDS), consideran que junto al CES-D, el PHQ-9 son los dos instrumentos de cribado que demuestran datos razonables para la detección de depresión en pacientes ambulatorios hispanoparlantes.
- Respecto a la versión española, un punto débil de la misma es la escasez de datos de validación psicométrica en muestras españolas de pacientes de AP.

3.1.5.2.- Cuestionario de Salud del Paciente–8 (PHQ-8)

Autor: K. Kroenke, T. Strine, R. L. Spitzer, J. B. Williams, J. T. Berry y A. H. Mokhad

Referencia: Kroenke, K., Strine, T., Spitzer, R. L., Williams, J. B., Berry, J. T., y Mokhad, A.H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114, 163-173.

Evalúa: Presencia y gravedad de síntomas de trastornos del estado de ánimo.

Nº de ítems: 8 ítems.

Administración: Autoaplicada

Descripción:

- El Cuestionario de Salud del Paciente-8 (*Patient Health Questionnaire-8*; PHQ-8) consiste en un cuestionario basado en los criterios DSM-IV (APA, 1994) y básicamente es igual a su predecesor, el PHQ-9 con la eliminación del último ítem (suicidio).
- Se estructura a través de cuatro categorías de respuesta que permite ver un cambio sintomático a lo largo del tiempo. La puntuación máxima es de 24. A diferencia del PHQ-9 el formato de respuesta está basado en la frecuencia medida en días. La equivalencia en este sentido sería:
 - 1 día = “nunca”
 - 2-6 días = “varios días”
 - 7-11 días = “más de la mitad de los días”
 - 11-14 días = “casi todos los días”

Corrección e interpretación:

- Los autores consideran dos formas alternativas de diagnosticar depresión:
 1. Algoritmo diagnóstico de la depresión basado en criterios DSM-IV. Se detecta depresión si puntúa en más de 5/8 ítems en la categoría “más de la mitad de los días” en las últimas dos semanas, siendo uno de los ítems sentimiento de tristeza o anhedonia. Otros tipos de depresión son diagnosticados si puntúa en 2, 3 ó 4 ítems, en las últimas dos semanas, siendo uno de los ítems sentimiento de tristeza o anhedonia.

2. Una puntuación total en el cuestionario mayor o igual a 10. Para medir la gravedad del episodio se consideran los criterios recogidos en la Tabla 3.8.

Tabla 3.8. Puntos de corte de PHQ-8 según Kroenke et al. (2009)

<i>Puntuación</i>	<i>Gravedad en depresión</i>
5-9	Ligera
10-14	Moderada
15-19	Moderada- Grave
20-24	Grave

Propiedades psicométricas:

- Validez de criterio: Según los autores, la curva ROC arroja una sensibilidad del 100% y una especificidad del 95 % para una puntuación mayor o igual a 10, y por tanto la probabilidad de que una persona con puntuación mayor o igual a 10 tenga depresión es 28 veces superior a que no lo tenga.

Desventajas:

- Los datos de validez de criterio tienen limitaciones porque se utilizó el algoritmo de diagnóstico del DSM IV para el PHQ-8 como criterio de referencia para los análisis psicométricos, en vez de utilizar una entrevista estructurada diagnóstica.

- No existen datos sobre muestras españolas.
- Los datos se han recabado vía telefónica.

3.1.6.- Escala Hospitalaria de Ansiedad y Depresión (HAD)

Autor: A. S. Zigmond y R. P. Snaith

Referencia: Zigmond, A. S., y Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361-370.

Evalúa: Detección de trastornos depresivos y ansiosos en el contexto hospitalario no psiquiátrico.

Nº de ítems: 14 ítems.

Administración: Autoaplicada

Descripción:

- La HAD fue creada con la intención de obtener un instrumento de detección de trastornos depresivos y ansiosos en el marco de los servicios hospitalarios no psiquiátricos que cumpliera las siguientes condiciones:
 - o No estar contaminado por la sintomatología de la patología física que presenta el paciente.
 - o Diferenciar claramente entre los trastornos de tipo depresivo y los de tipo ansioso.
- Se recomienda su uso como cribado; en el caso de que puntúe como positivo la evaluación ha de complementarse con otras escalas más específicas y heteroaplicadas (p. ej., la escala de depresión de Hamilton).

- La HAD está constituida por 14 ítems. Los ítems se agrupan en dos subescalas, cada una de ellas con 7 ítems:
 - o Subescala de ansiedad: está centrada en sus manifestaciones psíquicas (tensión, nerviosismo, aprensión, preocupación, inquietud, nervios en el estómago, angustia). Son los ítems impares.
 - o Subescala de depresión: centrada en la anhedonia (disfrute, risa, alegría, torpeza, interés por el aspecto personal, ilusión). Son los ítems pares.
- El paciente ha de contestar a cada ítem utilizando una escala de tipo *Likert* que unas veces hace referencia a la intensidad del síntoma y otras veces a la frecuencia de presentación. En cualquier caso, las escalas oscilan entre 0 (nunca) y 3 (casi todo el día, muy intenso).
- El marco de referencia temporal es la semana previa.
- Está adaptada por Snaith, en colaboración con Bulbena y Berrios en 1986 y validada al español por Tejero, Guimerá, Farré y Peri (1986) y Herrero et al. (2003). Existe otra adaptación española de la escala a cargo de Caro e Ibáñez (1992).

Corrección e interpretación:

- Para cada subescala la puntuación se obtiene sumando las puntuaciones en cada uno de los ítems que la conforman. El rango de puntuación es de 0-21 para cada subescala, y de 0-42 para la puntuación global.
- En ambas subescalas la puntuación obtenida se interpreta de acuerdo a los criterios que se presentan en la Tabla 3.9.

Tabla 3.9. Puntos de corte de la HAD según Caro e Ibáñez (1992)

<i>Puntuación</i>	<i>Gravedad en depresión/ansiedad</i>
0-2	Ausencia de depresión o ansiedad
8-10	Caso probable de depresión o ansiedad
11-21	Caso de ansiedad o depresión

Propiedades psicométricas:

- Fiabilidad: En la adaptación españolas se ha encontrado un $\alpha = 0,90$ para la puntuación global de ambas subescalas, $\alpha = 0,84$ para la subescala depresión, y $\alpha = 0,85$ para la subescala de ansiedad (Herrero et al., 2003). Otros estudios señalan un $\alpha = 0,86$ tanto para la escala de depresión como de ansiedad (Quintana et al., 2003). Terol et al. (2007), en una muestra no clínica (estudiantes universitarios y sus familiares), encuentran una consistencia interna de $\alpha = 0,77$ para la ansiedad y $\alpha = 0,71$ para la depresión, datos estos últimos algo inferiores a la media que aparece en estudios internacionales que se sitúan en un $\alpha = 0,82$ para la subescala de ansiedad y 0,83 para la de depresión (Bjelland et al., 2002). En cuanto a la fiabilidad test-retest es muy satisfactoria, por encima de 0,85 (Quintana et al., 2003). Terol et al. (2007) encuentran una correlación tes-retest de 0,77 para la ansiedad y de 0,74 para la depresión en una muestra de estudiantes.
- Validez de criterio: Un reciente metanálisis de 25 estudios encuentra que para el trastorno depresivo mayor, un punto de corte de ≥ 8 conlleva una

sensibilidad de 0,82 (95% IC: 0,73–0,89) y una especificidad de 0,74 (95% IC = 0,60–0,84) (Brennan, Worrall-Davies, McMillan, Gilbody y House, 2010). Si se aumenta el punto de corte a ≥ 11 encuentran una sensibilidad de 0,56 (95% I.C., .40–.71) y una especificidad de 0,92 (95% I.C. 0.79–0.97) (Brennan et al., 2010). En una muestra de estudiantes españoles, respecto a la depresión, el punto de corte 12 del total del HAD presenta un 0,80 de sensibilidad y un 0,73 de especificidad (Terol et al., 2007). En la revisión de Bjelland et al. (2002) se analizaron un total de 24 artículos, y los autores concluyen que en términos generales un punto de corte a partir de 8 tanto para la subescala de ansiedad como de depresión obtiene índices de sensibilidad y especificidad de aproximadamente 0,80. En el caso de AP encuentran dos estudios. En el primero de Wilkinson y Barczak (1988) con una muestra $N = 100$ y considerando los criterios DSM-III, se informa de que un punto óptimo a partir de 8 obtiene un área bajo la curva ROC de 0,96. El segundo, realizado por Rufaie y Absood (1995) con una muestra de $N = 217$ y utilizando la entrevista CIS (*Clinical Interview Schedule*) como criterio de referencia, se encuentra un área bajo la curva ROC de 0,86 (para las dos subescalas de ansiedad y depresión). En este último estudio se consideran como óptimos un punto de corte mayor de 9 para la subescala de ansiedad de la HADS, con una sensibilidad de 0,66 y especificidad de 0,93, y un punto de corte mayor de 7 para la subescala de depresión de la HAD, con una sensibilidad de 0,66 y una especificidad de 0,97 (citado en Bjelland et al., 2002).

- Validez factorial: Solución bidimensional, aunque las estructuras no coinciden exactamente con las escalas de ansiedad y depresión (Quintana

et al., 2003). En el metanálisis de Bjelland et al. (2000) se apunta que en 11 de los 19 artículos revisados en cuanto a su estructura factorial, habría consistencia en obtener una solución bidimensional aproximadamente coincidente con cada una de las subescalas de la HAD.

- Validez convergente y discriminante: En los últimos años, muchas han sido las críticas sobre la validez discriminante de la HAD. Bjelland et al. (2002) en su metanálisis encuentran 24 artículos que informan sobre la correlación (medida a través del coeficiente de Pearson) de las dos subescalas de ansiedad y depresión de la HAD con diferentes instrumentos de cribado. Como puede observarse en la Tabla 3.10, las diferencias entre los índices es baja. Tomada de forma conjunta, es decir, utilizando la puntuación total de la escala (HAD-T), parece que más bien detectaría un trastorno afectivo global, lo cual se ve refrendado por su validez convergente y discriminante aceptable con las diferentes escalas, en pocos casos excelente. En esta línea de considerar la puntuación total más que las dos subescalas, consideración cada vez con más apoyo empírico, Sultan (2010) obtiene mejor validez de criterio en una muestra de pacientes diabéticos al tomar la puntuación total (suma de las dos subescalas) que si toma cada puntuación por separado.

Tabla 3.10. Coeficientes de validez convergente y discriminante de la HAD con otros instrumentos de cribado (Bjelland et al., 2002)

<i>Referencia</i>	<i>Cuestionario de comparación</i>	<i>HAD-A</i>	<i>HADS-D</i>	<i>HAD-T</i>
Beck et al. (1997)	BDI-PC		0,62	
Lisspers et al (1997)	BDI	0,64	0,71	0,73
Suarez-Mendoza et al. (1997)	BDI		0,83	
Savard et al. (1998)	BDI	0,68	0,70	
Tedman et al. (1997)	BDI	0,61	0,73	
Watson et al. (1995)	BDI		0,69	
Lewis and Wessely (1990)	GHQ-12			0,75
Caplan (1994)	GHQ-28	0,68	0,66	
Chandarana et al. (1987)	GHQ-28	0,50	0,50	
Elliot (1993)	STAI	0,64	0,52	
Herrmann et al. (1991)	STAI	0,66	0,59	
Lisspers et al. (1997)	STAI-S	0,64		0,68
	STAI-T	0,66		
Millar et al. (1995)	STAI-S	0,81		
Savard et al. (1998)	STAI-S	0,78	0,65	
Lepine et al. (1985)	MADRS		0,62	
Snaith and Taylor (1985)	MADRS	0,37	0,81	
Upadhyaya and Stanley (1993)	MADRS		0,80	
Aylard et al. (1987)	MADRS (ITEM 3)			0,77
	CAS			0,67
Snaith and Taylor (1985)	CAS	0,69	0,44	
Upadhyaya and Stanley (1993)	CAS	0,75		
Spinhoven and van der Does (1997)	SCL-90	0,49	0,69	
Watson et al. (1995)	SCL-90	0,73	0,67	
Lepine et al. (1985)	HAM-S, HAM-P	0,34		
	HAM-T	0,40-0,44		
Millar et al. (1995)	VAS	0,74		

Nota: BDI: Beck Depression Inventory; BDI-PC: Beck Depression Inventory for Primary Care; CAS: Clinical Anxiety Scale; HAD-A: subescala de ansiedad de la HAD; HAD-D: subescala de depression de la HAD; HAD-T: puntuación total de la HAD; HAMA-S: Hamilton Anxiety Scale—Ítems somáticos; HAMA-P: Hamilton Anxiety Scale—Ítems psíquicos; HAMA-T: Hamilton Anxiety Scale—Escala total; MADRS: Montgomery-Asberg Depression Rating Scale; SCL-90: Symptom Checklist 90 Scale; STAI-S: subescala de Ansiedad Estado del Spielberger State-Trait Anxiety Inventory; STAI-T: subescala de Ansiedad Rasgo del Spielberger State-Trait Anxiety Inventory; VAS: Escala Visual Analógica.

Ventajas y desventajas:

- A tenor de los resultados encontrados, parece que es un instrumento con buenas propiedades psicométricas, lo que hace que sea útil en población general.
- Su brevedad también es un punto fuerte.
- Como puntos débiles señalar la baja discriminación entre ansiedad y depresión que antes se apuntaba.
- Otro aspecto a mejorar es que apenas introduce ningún síntoma somático. Aunque a partir de los datos encontrados en cuanto a sensibilidad y especificidad parece que el hecho de introducir síntomas somáticos no mejoraría la detección de los trastornos depresivos.

3.1.7. Inventario de Depresión de Beck-II (BDI II)

Autor: A. T. Beck, R. A. Steer y G. K. Brown.

Referencia: Beck, A. T., Steer, R. A., y Brown, G. K. (1996). *BDI-II. Beck Depression Inventory-Second Edition. Manual*. San Antonio, TX: The Psychological Corporation.

Evalúa: Presencia y gravedad de síntomas depresivos.

Nº de ítems: 21 ítems.

Administración: Autoaplicado.

Descripción:

- El BDI es, dentro de las escalas de depresión, la que mayor porcentaje de ítems cognitivos presenta, lo que está en consonancia con la teoría cognitiva de la depresión de Beck.
- El tiempo de aplicación es de aproximadamente 10 minutos.
- El formato de respuesta es mediante la elección de la frase que mejor describe al individuo de un grupo de cuatro frases alternativas para cada ítem ordenadas de menor a mayor gravedad (0-3).
- Diseñado para que diese cobertura a todos los criterios diagnósticos sintomáticos de los trastornos depresivos propuestos por el DSM-IV (APA, 1994).
- Respecto a la versión previa, el BDI-IA (Beck y Steer, 1978), se han remplazado 4 ítems (pérdida de peso, cambio en la imagen corporal, preocupación somática y dificultades en el trabajo) por otros 4 nuevos (agitación, sentimientos de inutilidad, dificultad de concentración y pérdida de energía). Además se han modificado los ítems para evaluar tanto decrementos como incrementos en apetito y en sueño. En total se han cambiado la redacción de 12 ítems.
- El marco temporal se ha incrementado con respecto a la anterior versión, siendo en el BDI-II de las dos últimas semanas.
- Ha sido adaptada y validada al castellano por el equipo de investigación de Sanz (Sanz, Navarro y Vázquez, 2003; Sanz, Perdigón y Vázquez, 2003; Sanz, García-Vera, Fortún, Espinosa y Vázquez, 2005).

Corrección e interpretación:

- La puntuación total se obtiene de la suma de cada ítem, pudiendo variar desde 0 a 63.
- En la Tabla 3.11 se muestran los puntos de corte recomendados por Beck et al. (1996) para distinguir distintos niveles de sintomatología depresiva.

Tabla 3.11. Puntos de corte del BDI-II según Beck, Steer y Brown (1996)

<i>Punto de corte</i>	<i>Gravedad de la depresión</i>
0–9	Depresión ausente
10–13	Depresión mínima
14–19	Depresión ligera
20–28	Depresión moderada
29–63	Depresión grave

- En una muestra española de adultos de población general, Sanz, Perdigón y Vázquez (2003) encontraron que una puntuación igual o menor de 9 en el BDI-II (la media de la población general española), podría ser un criterio para valorar mejoría o recuperación clínicamente significativa de un trastorno depresivo.

Propiedades psicométricas:

- Fiabilidad: En el manual original (Beck et al., 1996) se informa de un $\alpha = 0,92$ para pacientes con trastornos psicológicos y de un $\alpha = 0,93$ para estudiantes universitarios, lo que indica una excelente consistencia interna que mejora a la de su predecesor, el BDI-IA. El BDI-II también ha

demostrado excelente consistencia interna ($\alpha = 0,90$) en una muestra de afroamericanos con bajos recursos (Karen et al., 2005). Arnau, Meagher, Norris y Bramson (2001) han encontrado un $\alpha = 0,94$ en pacientes médicos de AP. Furukawa (2010) en su revisión considera que la fiabilidad medida a través del test-retest del BDI ha sido examinada solamente en muestras pequeñas, en muestras no clínicas y con índices inapropiados como, por ejemplo, la correlación de Pearson. A partir de estos datos, aunque escasos, considera (aún sin concretar el autor cuantificación) que el BDI II muestra una fiabilidad test-retest adecuada.

- Validez de contenido: Buena. Cubre 8/9 criterios del episodio depresivo mayor según criterios DSM-IV (APA, 1994), y 7/7 criterios de distimia según criterios DSM-IV (APA, 1994).
- Validez convergente: Satisfactoria con correlaciones moderadas o altas con la HDRS y la SDS de Zung. En su revisión, Furukawa (2010) también aprecia una aceptable validez concurrente con el HDRS de 0,71.
- Validez discriminante: Correlaciona 0,58 con el BAI, y 0,47 con la escala de Hamilton para la ansiedad (Furukawa, 2010).
- Validez de criterio: Muestra un índice kappa de concordancia = 0,59 con la entrevista diagnóstica DIS computerizada.
- Validez factorial: Los análisis factoriales suelen encontrar un factor cognitivo-afectivo y un factor somático-motivacional. Esta solución bifactorial ha sido identificada en pacientes psiquiátricos ambulatorios (Beck et al., 1996), pacientes de AP (Arnau et al., 2001), pacientes clínicamente deprimidos (Steer, Ball, Ranieri y Beck, 1999), pacientes geriátricos

deprimidos en régimen de ingreso (Steer, Rissmiller y Beck, 2000) y estudiantes de universidad (Beck et al., 1996b; Dozois, Dobson y Ahnberg, 1998). Del mismo modo, en la versión española del BDI-II en su aplicación a una muestra de pacientes médicos (hemodiálisis) se ha encontrado solución bifactorial parecida con dos factores primarios, uno cognitivo y otro afectivo, que a su vez conforman un factor de segundo orden denominado depresivo (Penley, Wiebe y Nwosu, 2003). Resultados similares aparecen en una muestra de pacientes afroamericanos de AP (Buttonm, Grothe y Jones, 2004). En términos generales, hay bastante acuerdo con la existencia de un factor general que mide depresión formado a su vez por dos factores altamente relacionados (Wiebe y Penley, 2005; Sanz y García-Vera, 2009). Quilty, Zhang y Bagby (2010), al realizar una revisión de la estructura latente del BDI-II, finalmente consideran estos autores un modelo trifactorial formado por un factor de depresión general, un factor somático y otro factor cognitivo (pese a la baja relevancia de las saturaciones de algunos ítems o no tener en cuenta el principio de parsimonia).

- Específicamente en una muestra de AP americana se han encontrado los siguientes índices de fiabilidad y validez para el BDI-II (Arnau, Meagher, Norris y Bramson, 2001):
 - Fiabilidad: Muy buena ($\alpha = 0,94$).
 - Validez de criterio: Área bajo la curva ROC igual a 0,96, lo que indica un rendimiento diagnóstico excelente como cribado del trastorno depresivo mayor. El punto de corte 18 presentaba los mejores índices diagnósticos, con una sensibilidad de 94%, una especificidad de 92%,

un valor predictivo positivo de 54% y un valor predictivo negativo de 99%.

- Correlación ítem-total: 0,54-0,74
- Los datos psicométricos en estudiantes universitarios españoles también se muestran aceptables: $\alpha = 0,89$ y la misma estructura factorial que en población general (Sanz, Navarro y Vázquez, 2003). Otros estudios también han mostrado similares propiedades psicométricas en adolescentes psiquiátricos ingresados (Osman, Kopper, Barrios, Gutierrez y Bagge, 2004).

Ventajas y limitaciones:

- En sus 50 años de uso, el BDI, en sus diferentes versiones, se ha convertido en uno de los instrumentos más ampliamente utilizados para detectar depresión en población general y para evaluar la gravedad de la depresión en pacientes clínicos adultos y adolescentes, y utilizado tanto por clínicos como por investigadores (Brantley, Dutton y Wood, 2004; Furukawa, 2010).
- El BDI II es un instrumento preciso y un eficiente instrumento de cribado para la depresión, mejorando la probabilidad de detectar sintomatología depresiva incluso en poblaciones minoritarias.

3.1.8.- Escala de Ansiedad y Depresión de Goldberg (EADG)

Autor: D. Goldberg

Referencia:

- Goldberg, D., Bridges, K., Duncan-Jones, P., et al. (1988). Detecting anxiety and depression in general medical settings. *British Medical Journal*, 97, 897-899.
- Goldberg, D., Bridges, K., Duncan-Jones, P., et al. (1989). Detección de la ansiedad y la depresión en el marco de la medicina general. *British Medical Journal (edición española)*, 4(2), 49-53.

Evalúa: Detección de trastornos depresivos y ansiosos en el contexto de la medicina general.

Nº de ítems: 18 ítems.

Administración: Heteroaplicada

Descripción:

- El Cuestionario de Goldberg o EADG (Escala de Ansiedad y Depresión de Goldberg) en la versión española fue desarrollado por este autor en 1988 a partir de una versión modificada de la *Psychiatric Assessment Schedule*, con la finalidad de lograr una entrevista de corta duración para ser utilizada por médicos no psiquiatras como instrumento de cribado.
- La versión en castellano ha sido validada por Montón et al. (1993).
- Consta de dos escalas, una de ansiedad y otra de depresión, con 9 ítems cada una, todos ellos de respuesta dicotómica (Sí/No).
- Marco temporal: 2 últimas semanas.

Corrección e interpretación:

- Cada una de las subescalas se estructura en 4 ítems iniciales de despistaje para determinar si es o no probable que exista un trastorno mental. Seguido de un segundo grupo de 5 ítems que se formulan sólo si se obtienen respuestas positivas a las preguntas de despistaje (2 o más en la subescala de ansiedad, 1 o más en la subescala de depresión).
- Se obtiene una puntuación independiente para cada escala, contabilizándose un punto por cada respuesta afirmativa. Los puntos de corte para detectar un caso de depresión o un caso de ansiedad se recogen en la Tabla 3.12.

Tabla 3.12. Puntos de corte para la EADG (Goldberg, 1989)

<i>Subescala de Depresión</i>	<i>Subescala de Ansiedad</i>
≥ 4	≥ 2

Propiedades psicométricas:

- Validez de criterio: En población geriátrica la validez predictiva de la subescala de ansiedad es pobre (Neal y Basldwin, 1994) y la capacidad discriminante de las dos subescalas es menor, por lo que se ha propuesto su uso como escala única, con un punto de corte ≥ 6 (Mackinnon et al 1994). Montón et al. (1993) encuentran en su estudio una sensibilidad del 83% y una especificidad del 82%.

Ventajas e inconvenientes:

- Su sencillez, unida a los buenos índices de sensibilidad y especificidad, a su capacidad discriminante entre ansiedad y depresión y a su capacidad de aportar información dimensional sobre gravedad, han hecho que esta escala haya sido recomendada como instrumento de cribado, tanto con fines asistenciales, como epidemiológicos o como guía de la entrevista clínica en el ámbito de la AP (Buitrago, Ciurana y Chocron, 1999).
- Sin embargo, las propiedades psicométricas del EADG han sido escasamente estudiadas y los estudios que se encuentran tienen deficiencias metodológicas, lo que dificulta una caracterización precisa del instrumento.

A modo de resumen, en la Tabla 3.13 se presentan las principales características de los instrumentos de cribado para la depresión más relevantes de los que se utilizan en España y sobre los cuales, en el presente capítulo, se han revisado sus principales índices psicométricos. Estos índices permiten tener una base de datos sobre la que comparar y poder evaluar de un modo más preciso los datos que se obtengan en la investigación empírica de esta tesis doctoral tanto con el BDI-FS como con los instrumentos que se utilizarán de manera complementaria para analizar la propiedades psicométricas del BDI-FS.

Tabla 3.13. Cuadro resumen de los principales instrumentos de cribado de depresión utilizados en España

<i>Test</i>	<i>Nº Items</i>	<i>Contenido</i>	<i>Formato respuesta</i>	<i>Espacio temporal</i>	<i>Puntuación</i>	<i>Punto de corte</i>	<i>Medida de gravedad</i>
BDI II (Beck, Steer y Brown, 1996)	21	Depresión	4 afirmaciones por ítem Autoaplicada	Últimas dos semanas	0-63	≥ 10	Si
BDI FS (Beck, Steer y Brown, 1999)	7				0-21	≥ 4	Si
CES-D (Radlof, 1977)	20, 10	Depresión	Escala Likert (4) Autoaplicada	Última semana	0-60	≥ 16	Si
PRIME-MD (Spitzer, Williams, Kroenke, et al., 1994)	25	Múltiples trastornos	Dicotómica (Sí/No) Autoaplicada	Último mes	0-2	≥ 1	No
PHQ (Spitzer, Kroenke, Williams, et al. 1999)	9, 8	Depresión	Dicotómica (Sí/No) Autoaplicada	Último mes	0-9 por diagnóstico o 0-27 por respuesta	≥ 5	Si
SDS (Zung, 1965)	20	Depresión	Escala Likert (4) Autoaplicada	Reciente-mente	0-80	≥ 50	Si
EADG (Goldberg, 1988)	18	Depresión y ansiedad	Dicotómica (Sí/No) Heteroaplicada	2 últimas semanas	0-9 en cada escala.	≥ 2 depresión ≥ 4 ansiedad	No
HDRS (Hamilton, 1960, 1967)	6, 7, 17, 21 y 25	Depresión	Escala Likert (4) Heteroaplicada	Momento actual, en algunos ítems 2 días previos	0-60	≥ 18	No
HADS Zigmond y Snaith (1983)	14	Depresión y ansiedad	Escala Likert (4) Autoaplicada	Última semana	0-21	≥ 11	Si

Capítulo 4:

El Inventario de Depresión de Beck y su versión para pacientes médicos: el BDI-FS

*“Los hombres no se perturban por causas de las cosas,
sino por la interpretación que de ellas hacen”*

Marco Aurelio Epicteto en “El Enchiridión”

4.1. El Inventario de Depresión de Beck (BDI)

A principios de los años 60 del siglo pasado, Aaron T. Beck, psiquiatra internacionalmente conocido por ser el padre de la terapia cognitiva para la depresión, creó junto a su equipo de investigación el Inventario de Depresión de Beck (*Beck Depression Inventory* o BDI; Beck, Ward, Mendelson, Mock y Erbaugh, 1961). El BDI se construyó “para medir las manifestaciones conductuales de la depresión” y “disponer de una evaluación cuantitativa de la intensidad de la depresión” (Beck et al., 1961, pp. 53 y 61).

A lo largo de sus 50 años de vida, el BDI ha sufrido varias revisiones y modificaciones y por esa razón para distinguir el BDI original de revisiones o versiones posteriores, a veces al original se le denomina BDI-I. A finales de los años 70 del siglo XX, Beck y sus colaboradores empezaron a trabajar en una versión modificada o corregida del BDI original y que actualmente se conoce como BDI-IA (BDI-I *amended* o corregido). Esta versión se publicó por primera vez en el clásico manual de terapia cognitiva de la depresión de Beck, Rush, Shaw y Emery (1979), aunque su patente fue registrada en 1978, por lo que a veces también se le denomina BDI de 1978. En los años 90 del siglo XX Beck y sus colaboradores realizaron una revisión profunda del BDI-IA para adecuar su contenido a los criterios diagnósticos de los trastornos depresivos del DSM-IV. Esta versión presentaba modificaciones tan sustanciales respecto al BDI-I y al BDI-IA, que sus autores la denominaron BDI-Segunda Edición o BDI-II, y fue publicada en 1996 (Beck, Steer y Brown, 1996). Durante todos estos años también se han creado versiones cortas del BDI-I, del BDI-IA y del BDI-II, de las cuales las más conocidas son el BDI-Forma Breve o BDI-SF (*BDI-Short Form*) de Beck y Beck (1972), una versión breve del BDI original de 1961

pensada para identificar rápidamente la depresión y estimar su gravedad en pacientes con enfermedades médicas, y el BDI para Atención Primaria o BDI-PC (*BDI for Primary Care*) el cual, actualmente, se denomina BDI-Cribado Rápido o BDI-FS (*BDI-Fast Screen*) (Beck, Steer y Brown, 2000), y que es un versión breve del BDI-II pensada también para evaluar la depresión en pacientes médicos y para detectar los pacientes médicos que sufren trastornos depresivos.

Teniendo en cuenta sus diferentes versiones y revisiones, actualmente se considera al BDI a nivel internacional como el test autoaplicado más utilizado en la práctica clínica y en la investigación para medir la sintomatología depresiva en pacientes con trastornos psicológicos y en personas de la población normal (Beck, Steer y Garbin, 1988; Piotrowski, 1996; Sanz, Navarro y Vázquez, 2003), de manera que existen traducciones o adaptaciones de las diferentes versiones del BDI (BDI, BDI-IA, BDI-II, BDI-SF, BDI-PC ó BDI-FS) a una gran multitud de idiomas: alemán, árabe, búlgaro, chino, danés, español, finlandés, flamenco, francés, griego, hebreo, hindú, italiano, islandés, japonés, noruego, sueco, turco, etc. (Sanz, 2010).

Consistentemente, también en España el BDI se ha convertido en el test de autoinforme más utilizado para evaluar la depresión tanto en la práctica clínica (Muñiz y Fernández-Hermida, 2000) como en investigación (Sanz, Navarro y Vázquez, 2003).

4.2. El BDI original (BDI-I) y su adaptación española

La versión original del BDI o BDI-I fue publicada en 1961. Mientras el psicoanálisis imperaba en la época, Beck trataba de construir un instrumento que recabara de manera comprehensiva las creencias y sentimientos de los pacientes depresivos independientemente de las diferentes teorizaciones.

Obtuvo así un conjunto de 21 ítems compuestos por grupos de frases o alternativas de respuesta sobre un mismo síntoma depresivo. De esos 21 ítems, 12 contienen cuatro frases u opciones de respuesta, ocho cinco opciones y un ítem seis opciones, estando las opciones o frases graduadas en función de la gravedad del síntoma depresivo que reflejan. El marco temporal por el que se pregunta a la persona evaluada es el momento en el que se cumplimenta el test.

Independientemente del número de alternativas de los ítems, la puntuación de todos ellos varía de cero a tres. Así, en algunos ítems se puede escoger entre dos frases que obtienen la misma puntuación (p. ej., el ítem que pregunta sobre el estado de ánimo). En el caso de que el sujeto elija más de una alternativa en un ítem dado, se considera sólo la puntuación de la frase elegida de mayor gravedad. La puntuación total (0-63) se obtiene a partir del sumatorio de las puntuaciones de todos los ítems. Los puntos de corte que se establecieron para el BDI-I se recogen en la Tabla 4.1.

En España, Conde, Esteban y Useros (1976) realizaron la adaptación al castellano del BDI-I. Cabe señalar alguna de las limitaciones de esta adaptación española como, por ejemplo, la presentación no graduada de las diferentes afirmaciones dentro de un mismo ítem, lo que dificulta las

comparaciones con la versión original en la cual todas las afirmaciones se presentan graduadas, empezando con la que refleja la menor gravedad. Otro aspecto relevante de la adaptación española de Conde et al. (1976) es la eliminación de los ítems referidos al sentimiento de castigo y a la pobre imagen corporal, dejando a la versión española con 19 ítems y, por tanto, dificultando también las comparaciones con las investigaciones internacionales (Sanz y Vazquez, 1998).

Tabla 4.1. Puntos de corte del BDI-I (Beck, Epstein, Brown y Steer, 1988)

<i>Punto de corte</i>	<i>Gravedad de la depresión</i>
0–9	Depresión ausente
10–18	Depresión leve
19–29	Depresión moderada
30–63	Depresión grave

Como se dijo en el epígrafe anterior, de esta versión del BDI existe una versión breve creada por Beck y Beck (1972) llamada *BDI Short Form* (BDI-SF) y diseñada para identificar rápidamente la depresión y estimar su gravedad en pacientes con enfermedades médicas. Está formada por los 13 ítems del BDI-I que mejor correlacionaban con la puntuación total del BDI-I y con las valoraciones de la gravedad de la depresión realizada con el criterio clínico. Su estudio ha sido bastante prolífico en diferentes muestras, obteniendo interesantes propiedades psicométricas. De este modo, se han encontrado índices adecuados de consistencia interna (Gould, 1982; Knight, 1984; Leahy,

1992; Reynolds y Gould, 1981; Van Hemert, Van de Vijver y Poortinga, 2002), validez factorial (Gould, 1982; Leahy, 1992; Volk, Pace y Parchman, 1993), y validez convergente con otros instrumentos que miden sintomatología depresiva (Beck, Rial y Rickels, 1974; Cathebras, Mosnier, Levy, Bouchou y Rousset, 1994; Reynolds y Gould, 1981; Stukenberg, Dura y Kiecolt-Glaser, 1990) (todas las referencias psicométricas citadas en Sanz y García-Vera, 2007). En cuanto a su validez de criterio, Chochinov, Wilson, Enns y Lander, (1997) encuentran en un grupo de 197 pacientes terminales que una puntuación mayor o igual a ocho en el BDI-SF obtiene una sensibilidad de 0,79 y una especificidad de 0,71. Cabe señalar la elevada tasa de falsos positivos (29%), lo cual es importante según los objetivos que se planteen en la muestra a estudiar.

En esta línea y como desventaja del BDI-SF aparece, al igual que en el BDI-I del que surge, la referencia temporal “actual”, la cual se aleja bastante de las dos semanas necesarias actualmente para el diagnóstico del episodio depresivo mayor según criterios DSM-IV (APA, 1994).

4.3. El BDI-I corregido (BDI-IA) o BDI de 1978 y su adaptación española

En los años 70 del siglo pasado, Beck y sus colaboradores crean una versión revisada o corregida del BDI-I, patentada en 1978, publicada por primera vez un año después y conocida como BDI-IA o BDI de 1978 (Beck, Rush, Shaw y Emery, 1979; Beck y Steer, 1993). El objetivo de esta revisión fue tratar de facilitar la comprensibilidad de los 21 ítems del BDI original, lo cual

al tratarse de un autoinforme es un aspecto que cobra especial relevancia. Dichos cambios se concretaron en la eliminación de las dobles negaciones así como de las opciones dentro de un mismo ítem que superaban el número de cuatro, de forma que al final se modificaron un total de 15 ítems del BDI-I. Además, el marco temporal para esta versión corregida se amplía a “la semana anterior teniendo en cuenta el día de hoy”.

Como ocurría en el BDI-I, la puntuación en cada ítem del BDI-IA varía de 0 a 3, estando las frases u opciones graduadas según su gravedad. La corrección se hace a través del sumatorio de sus 21 ítems. La pérdida de peso (ítem 19) sólo se valora si el sujeto indica no estar bajo dieta para adelgazar. En el caso de que lo esté, se otorga una puntuación de 0 en el ítem.

Los puntos de corte variaron ligeramente respecto a su predecesor, siendo necesario un descenso de la puntuación para los criterios de depresión moderada y grave. En la Tabla 4.2 se recogen los puntos de cortes propuestos por Beck y Steer (1993) para el BDI-IA.

Tabla 4.2 Puntos de corte del BDI-IA según Beck y Steer (1993)

<i>Punto de corte</i>	<i>Gravedad</i>
0–9	Depresión ausente
10–16	Depresión leve
17–29	Depresión moderada
30–63	Depresión grave

En España existe una adaptación de la versión revisada de 1978 (Sanz y Vázquez, 1998; Vázquez y Sanz, 1991, 1997, 1999), que presenta buenos índices de fiabilidad y validez en consonancia con los que presenta el BDI-IA original en lengua inglesa. Por ejemplo, Lasa, Ayuso-Mateos, Vázquez-Barquero, Díez-Manrique y Dowrick (2000) encontraron, en una muestra aleatoria de 1250 personas de Santander, que las puntuaciones de 12 ó 13 obtienen un 98% de eficiencia diagnóstica en trastornos depresivos, lo que sustentaría su utilidad como instrumento para detectar trastornos depresivos en la población general española. Por su parte, Vázquez y Sanz (1997) encontraron en una muestra de 338 pacientes psicopatológicos españoles, que la puntuación de 21 obtenía un 76% de eficiencia diagnóstica a la hora de detectar trastornos depresivos en la población clínica española con trastornos mentales.

A pesar de las buenas propiedades psicométricas que tienen el BDI-IA tanto en su versión original como en su adaptación española, esta versión tiene sus limitaciones. Especialmente cabe resaltar que sólo medía seis de los nueve criterios requeridos para el episodio depresivo mayor según el DSM-III (APA, 1980). Si tomamos la definición de episodio depresivo mayor más consensuada hoy en día, la ofrecida por el DSM-IV (APA, 1994), el BDI-IA no cubre el criterio sintomático de enlentecimiento o agitación psicomotores, y sólo refleja de manera parcial (recogiendo sólo los aspectos deficitarios pero no de exceso) otros dos: problemas de sueño y problemas de apetito/peso (Vázquez y Sanz, 1997). Por otro lado, el BDI-IA cubre síntomas (p. ej., hipocondría) que no son criterios diagnósticos del DSM-IV para el episodio depresivo mayor. Otro inconveniente a tener en cuenta es que el BDI-IA no permite evaluar la

presencia de sintomatología depresiva durante el período mínimo necesario para el diagnóstico de episodio depresivo mayor (dos semanas en el DSM-IV frente al marco temporal de una semana utilizado por el BDI-IA).

Actualmente, el BDI-IA es uno de los instrumentos más utilizados en clínica psiquiátrica y psicológica para evaluar la gravedad de la depresión en pacientes con trastornos mentales y para detectar posibles depresiones en población general (Archer, Maruish, Imhof y Piotrowski, 1991; Piotrowski y Keller, 1992; Piotrowski, Sherry y Keller, 1985; citados en Beck, Steer y Brown, 2000). De hecho, pese a la superioridad del BDI-II respecto a su predecesor el BDI-IA, tanto a nivel de consistencia interna, validez factorial (Sanz y García-Vera, 2009) o validez de contenido (el BDI-II cubre en mayor medida los criterios del DSM IV para la depresión), el BDI-IA es actualmente el cuestionario más conocido y utilizado en España tanto a nivel de clínica como de investigación.

Con el fin de optimizar el tiempo de aplicación del instrumento, con apenas menoscabo de sus bondades psicométricas, de manera que se puedan usar conjuntamente otros instrumentos para patologías comórbidas, aparecen dos versiones breves del BDI-IA (Sanz y García-Vera, 2007).

La primera versión corta se gesta a partir del análisis factorial del BDI-IA y parte de la idea de Plumb y Holland (1977) sobre la posibilidad de eliminar los síntomas somáticos del BDI-I (1961) para medir la depresión y su gravedad en personas con enfermedad física. Beck y Steer (1993) consideraron el uso de la subescala cognitiva- afectiva (ítems 1-13) para la medición de la depresión y su gravedad en pacientes con trastorno de abuso de sustancias o enfermedades médicas. De esta forma, al eliminar el componente somático (ítems 14 a 21)

evitaban el solapamiento de síntomas físicos. Así, los ítems que conforman esta versión de versión abreviada del **BDI-IA basada en la subescala cognitivo afectiva** (a partir de ahora BDI-IA-SCA) serían: tristeza, pesimismo, sentimiento de castigo, autodecepción, autoinculpación, ideas de suicidio, llanto, irritabilidad, desinterés social, e indecisión.

En relación con su validez de criterio, Furlanetto, Mendlowicz y Romildo Bueno (2005) encuentran una buena capacidad del BDI-IA-SCA para diferenciar deprimidos y no deprimidos en una muestra de pacientes hospitalizados con enfermedades médicas.

Sanz y García Vera (2007) hallaron a través de un análisis de regresión de la puntuación total del BDI-IA una ecuación que permitía estimar los puntos de corte y diferentes rangos de gravedad en la depresión en el BDI-IA-SCA. Las puntuaciones propuestas se muestran en la Tabla 4.3.

Tabla 4.3. Puntos de corte del BDI-IA-SCA según Sanz y García-Vera (2007)

<i>Puntuación</i>	<i>Gravedad de la depresión</i>
0-6	Mínima
7-11	Ligera
12-20	Moderada
21-39	Grave

En España, Sanz y García Vera (2007) han estudiado las propiedades psicométricas de las versiones breves del BDI-IA en tres muestras: 338 personas con psicopatología, 445 personas procedentes de población general y 1393 estudiantes universitarios. El análisis de la consistencia interna encontró datos muy positivos para el BDI-IA SCA, encontrándose un α de 0,89, 0,76 y 0,79 para las muestras de pacientes, de la población general y de estudiantes, respectivamente. Todos estos datos son considerados como aceptables según los estándares de Cicchetti (1994), dado que superan el valor de 0,70 considerado como aceptable para la consistencia interna. En todos los casos los análisis factoriales a través de ejes principales revelaban que el BDI-IA-SCA medía una dimensión general de depresión formada por una solución bifactorial que comprendía un factor de corte cognitivo y otro motivacional-afectivo. En cuanto a la validez de criterio encontraron que un punto de corte de 15 alcanzaba un grado bueno tanto de sensibilidad (72,9%) como de especificidad (80%). Su validez de contenido es menor a la del BDI-IA-SF al no considerar, por su propios objetivos los síntomas somáticos, ni tampoco criterio de deterioro o disfuncionalidad, de tal forma que cubre 5/9 síntomas del episodio depresivo mayor y 4/7 de la distimia según el DSM-IV (APA, 1994).

La segunda versión corta del BDI-IA viene propuesta por Bennett et al. (1997), los cuales sugirieron utilizar una versión breve del BDI-IA que incluyera los ítems de éste que se correspondían con los de la versión del BDI-SF (Beck y Beck, 1972). De tal forma que los ítems que conforman esta versión, denominada **BDI-IA-SF**, serían: tristeza, pesimismo, sentimiento de fracaso, insatisfacción, culpabilidad, autodecepción, ideas de suicidio, desinterés social, indecisión, pobre imagen corporal, dificultad laboral, fatiga, y pérdida de apetito.

Su repercusión no ha sido muy extensa, pese a su buena validez de criterio a la hora de diferenciar entre depresivos y no depresivos en pacientes médicos. Love, Grabsch, Clarke, Bloch y Kissane (2004) en un estudio de 227 mujeres con metástasis de cáncer de mama y con un punto de corte de cuatro en el BDI-IA-SF, obtuvieron una sensibilidad satisfactoria de 0,84 en detrimento de una especificidad de 0,63. En este estudio, pese a citar el uso de la versión corta del BDI-I (BDI-SF), si se tiene en cuenta la referencia temporal que se cita en la investigación (periodo de 7 días), parece que podrían estar haciendo uso de la versión posterior (BDI-IA-SF).

Del mismo modo que en la versión breve anteriormente comentada, Sanz y García Vera (2007) encontraron una ecuación que permitía estimar las puntuaciones de corte para distintos niveles de gravedad en el BDI-IA-SF (véase la Tabla 4.4).

Tabla 4.4. Puntos de corte del BDI-IA-SF según Sanz y García-Vera (2007)

<i>Puntuación</i>	<i>Gravedad de la depresión</i>
0-6	Mínima
1-11	Ligera
12-20	Moderada
21-39	Grave

En el estudio de Sanz y García Vera (2007), el análisis de la consistencia interna encontró datos muy positivos para el BDI-IA-SF, hallándose un α de 0,79, 0,80, 0,79 para las muestras de pacientes, de la población general y de estudiantes, respectivamente. Todos estos datos son considerados como aceptables según los estándares de Cicchetti (1994), dado que superaban el valor de 0,70 considerado como aceptable para la consistencia interna. Los datos sobre validez factorial también arrojaban una solución bifactorial compuesta por dos factores (cognitivo y somático) altamente correlacionados entre sí, lo que del mismo modo que en la otra versión breve y en general en el resto de versiones del BDI, hacen considerar la medida de una dimensión general de depresión. En este mismo estudio un punto de corte de 14 sería el valor óptimo que alcanzaría unos niveles de sensibilidad del 72,9% y de especificidad del 76,9%, lo que supondría una buena validez de criterio, es decir, una buena capacidad para discriminar entre pacientes deprimidos y no deprimidos. El área bajo la curva ROC encontrada fue de 0,82 algo superior al 0,81 hallado para la versión completa (BDI-IA). En cuanto a su validez de contenido, el BDI-IA-SF cubre 7/9 síntomas para el episodio depresivo mayor y 5/7 para la distimia, además cubre el criterio C según el DSM-IV (APA, 1994) a través de un ítem que evalúa la dificultad laboral.

Finalmente, pese a las buenas propiedades psicométricas y conseguir los objetivos pretendidos, estas dos versiones apenas han tenido éxito tanto a nivel nacional como internacional en cuanto a estudios más profundos en investigación o difusión en el ámbito clínico.

4.4. La segunda edición del BDI (BDI-II) y su adaptación española

En 1996 se publica la última revisión del BDI o segunda edición del BDI, conocida como BDI-II (Beck, Steer y Brown, 1996). Al igual que sus antecesores, se trata de un instrumento autoaplicado de 21 ítems cuyo objetivo es medir la sintomatología depresiva, así como cuantificar su gravedad, pero incluye importantes modificaciones sobre su predecesor, el BDI-IA (Beck y Steer, 1993), con el fin de recoger los criterios diagnósticos para el episodio depresivo mayor según el DSM-IV (APA, 1994).

Su adaptación al castellano ha sido realizada por el equipo de investigación de Sanz (Sanz, Navarro y Vázquez, 2003; Sanz, Perdigón y Vázquez, 2003; Sanz, García-Vera, Espinosa, Fortún y Vázquez, 2005).

En el BDI-II prácticamente todos los ítems del BDI-IA han sido revisados. Se reemplazaron cuatro ítems del BDI-IA (pérdida de peso, cambio en la imagen corporal, preocupación somática y dificultades en el trabajo) por otros tantos nuevos (agitación, sentimientos de inutilidad, dificultad de concentración y pérdida de energía). Se modificaron los ítems de pérdida de apetito e insomnio para que éstos pudieran evaluar tanto los decrementos como los incrementos en apetito y sueño, y se introdujeron modificaciones en la redacción de otros doce ítems. Además, el marco temporal ha aumentado a dos semanas con el fin de acoger el tiempo necesario para diagnosticar el episodio depresivo mayor según el DSM-IV (APA, 1994).

Existe gran cantidad de estudios psicométricos que avalan las buenas propiedades psicométricas del BDI-II, tanto respecto a la **fiabilidad**,

consiguiendo niveles de consistencia interna entre adecuados y excelentes, superiores al 0,70 y en la mayoría de los estudios por encima del 0,80 (Beck, Steer y Brown, 1996; Beck, Steer, Ball et al., 1996; Steer et al., 1997; 1998; Acton et al., 2001; Sprinkle et al., 2001), como en cuanto a la **validez** BDI-II en diferentes muestras: pacientes psiquiátricos (Beck, Steer y Brown, 1996), pacientes con trastornos depresivos (Steer, Ball, Ranieri y Beck, 1999), adolescentes (Coelho, Martins y Barros, 2002), drogodependientes (Buckley, Parker y Heggie, 2001), ancianos (Jefferson, Powers y Pope, 2001), pacientes de atención primaria (Arnau, Meagher, Norris y Bramson, 2001) y estudiantes universitarios (Al Musawi, 2001; Beck, Steer y Brown, 1996; Dozois, Dobson y Ahnberg, 1998; O'Hara, Sprinkle y Ricci, 1998; Osman, Downs, Barrios, Kopper, Gutierrez y Chiro, 1997; Schulenberg y Yutrzenka, 2001; Steer y Clark, 1997; Whisman, Perez y Ramel, 2000) (referencias citadas en Sanz, Navarro y Vázquez, 2003).

En resumen, el BDI-II presenta, respecto a su predecesor, el BDI-IA, índices de fiabilidad y validez tan buenos o incluso superiores (Beck, Steer, Ball y Ranieri, 1996; Beck, Steer y Brown, 1996; Sanz, Navarro y Vázquez, 2003).

En cuanto a la versión española del BDI-II, sus propiedades psicométricas han sido estudiadas en muestras de estudiantes universitarios (Sanz, Navarro y Vazquez, 2003), en muestras de adultos procedentes de la población general (Sanz, Perdigon y Vazquez, 2003) y en muestras clínicas (Sanz, García-Vera, Espinosa, Fortún y Vázquez, 2005), obteniendo en todas ellas índices psicométricos adecuados que justifican su uso incluso por encima de las versiones anteriores (Sanz y García Vera, 2009).

En los últimos años, Sanz, García-Vera, Fortún y Espinosa (2005) han creado una versión breve del BDI-II (BDI-II-SF) que incluye los 11 ítems del BDI-II que mejor se comportan en términos de fiabilidad y validez en una muestra española de pacientes con trastornos psicológicos. Serán necesarios más estudios de esta forma breve para ver su funcionamiento en diferentes muestras.

4.5. El Inventario de Depresión de Beck-Cribado Rápido (BDI-FS)

A partir de la última versión del BDI, el BDI-II, y siguiendo la misma línea que habían hecho con el BDI-IA (1978) y su forma breve, el BDI-IA-SF(1993), Beck, Steer y Brown (2000) hipotetizaron que creando un instrumento formado por los ítems de carácter no somáticos del BDI-II, éste podría ser útil para detectar a aquellos pacientes depresivos que estarían atribuyendo sus síntomas somáticos o conductuales a factores biológicos, médicos, de abuso de alcohol o de sustancias. De este modo, se podría identificar a aquellos pacientes que pudieran estar padeciendo un cuadro depresivo y proceder a una evaluación más exhaustiva con el fin de realizar el diagnóstico preciso.

Así, con los ítems cognitivos y afectivos del BDI-II se construye el BDI-FS (*BDI-Fast Screen* o BDI-Cribado Rápido para pacientes médicos), inicialmente conocido como *BDI for Primary Care* (BDI-PC) o BDI para Atención Primaria (Beck, Guth, Steer y Ball, 1997). El BDI-FS surge en respuesta a la necesidad de facilitar el cribado y detección de la depresión en el marco de la AP, y con el objetivo de disponer de un instrumento breve (se tarda menos de 5

minutos, excepto en pacientes obsesivos o muy deprimidos), sencillo y rápido, con tan buenas propiedades psicométricas como sus antecesores, que permita la identificación de la depresión y su evaluación en pacientes de AP.

El BDI-FS es un instrumento de siete ítems autoaplicados (aunque como todas las versiones del BDI se puede administrar de forma heteroaplicada). Los ítems de *tristeza* y *anhedonia* fueron incluidos por su requerimiento según el DSM-IV (APA, 1996) para realizar el diagnóstico de depresión. *Pensamiento o deseo de suicidio* fue incorporado por su relevancia como indicador de riesgo suicida en el paciente depresivo. *Pesimismo, insatisfacción con uno mismo, sentimiento de fracaso y autocrítica* fueron escogidos por su relevancia factorial (saturaciones factoriales $\geq 0,35$) en el factor cognitivo del BDI-II (Beck, Steer y Brown, 1996). La puntuación total del BDI-FS se obtiene sumando la mayor de las puntuaciones en cada uno de los siete ítems. Cada ítem puntúa de 0 a 3 en función de la frase elegida como respuesta. La puntuación máxima es 21. El BDI-FS se contextualiza en las dos últimas semanas incluyendo el día de hoy, con lo que se adscribe al criterio temporal necesario para el diagnóstico del trastorno depresivo según el DSM-IV (1994). No correlaciona de manera significativa con sexo, etnia o edad.

Teniendo en cuenta una puntuación media de ocho puntos en las puntuaciones de los pacientes procedentes de dispositivos médicos y una desviación típica de cuatro puntos, los autores recomiendan los siguientes puntos de corte que aparecen en la Tabla 4.5 para valorar el grado de depresión.

Tabla 4.5. Puntos de corte del BDI-FS según Beck et al. (2000)

<i>Puntuación</i>	<i>Gravedad de la depresión</i>
0-3	Mínima
4-8	Ligera
9-12	Moderada
13-21	Grave

En relación con la monitorización de los cambios en sintomatología depresiva debidos, por ejemplo, a la aplicación de un tratamiento para la depresión, los autores del BDI-FS sugieren el uso de la desviación típica de la muestra objeto de estudio como criterio para considerar que ha habido un cambio relevante en la sintomatología depresiva entre un momento (p. ej., pretratamiento) y otro (p. ej., postratamiento).

Además, los autores recomiendan el uso del BDI-FS por personas profesionales sanitarios con suficiente experiencia, que conozcan el manejo del paciente depresivo y que estén familiarizados con el uso de instrumentos de evaluación (Beck et al., 2000).

Como puede verse en la Tabla 4.6, el BDI-FS aporta unos índices psicométricos muy interesantes. Esta versión aporta brevedad con apenas diferencia en cuanto a la calidad de sus índices psicométricos. De esta forma,

su fiabilidad medida a través de consistencia varía desde un α de Cronbach de 0,83 a 0,88, y en todos los casos se considera aceptable y es comparable con el $\alpha = 0,83$ y 0,87 del BDI-IA y del BDI-II, respectivamente. En cuanto a su validez de criterio, los autores (Beck et al., 2000) consideran que, en términos generales y teniendo en cuenta que el punto de corte óptimo varía a través de las muestras, un punto de corte entre tres y cinco es recomendable para conseguir la máxima eficiencia en poblaciones médicas donde la prevalencia para los trastornos depresivos oscila desde un 5-40%. De manera más concreta, Beck et al. (1997) recomiendan una puntuación ≥ 4 para obtener la máxima eficiencia en muestras de pacientes con enfermedad médica. En este sentido, la capacidad de discriminación entre pacientes depresivos y no depresivos medido a través del área de la curva ROC varía desde un 0,85 hasta un 0,99, lo cual, confirma la buena capacidad de detección del BDI-FS. Su eficiencia oscila entre 82-98%. Los datos sobre sensibilidad también son muy positivos variando desde 74% a 97%. Además, destaca por su capacidad para descartar pacientes no depresivos, de manera que su especificidad va desde 67% hasta 99%.

Tabla 4.6. Resumen de las propiedades psicométricas del BDI-FS en la literatura científica

Referencia	N	País	Muestra	% de mujeres	Puntuación media (DT)	Criterio oro	Fiabilidad (<i>alfa</i>)	Área bajo la curva ROC	Sensibilidad	Especificidad	Eficiencia	Punto de corte
Steer et al. (1999)	120	EEUU	Atención primaria	50%	2,18 (2,96)	PRIME-MD	0,85	0,99	97%	99%	98%	>4
Beck, Guth, Steer y Ball (1997)	50	EEUU	Ingreso en medicina interna	60%	5,76 (4,46)	PRIME-MD	0,86	0,92	82%	82%	82%	>4
Poole et al. (2009)	1227	EEUU	Dolor crónico	62%	7,11 (4,30)	BDI-II	0,83	0,94 0,93	81% 74%	92% 93%		>4 >5
Krefetz et al. (2004)	63	EEUU	Pacientes con VIH con dolor crónico	30%	7,69 (3,92)	PRIME-MD	0,84	0,86	90%	74%	81%	>4
Orive et al. (2010)	46	España	Clínica médica	63%		PRIME-MD	0,86	0,85	82%	72%		
Wilheim et al. (2004)	213	Australia	Psiquiatría de enlace	55,2%		Criterio Médico		0,84-0,89	75-91%	62-72%		>4

Referencia	N	País	Muestra	% de mujeres	Puntuación media (DT)	Criterio oro	Fiabilidad (<i>alfa</i>)	Área bajo la curva ROC	Sensibilidad	Especificidad	Eficiencia	Punto de corte
Sdheinthal et al. (2001)	75	EEUU	Geriatría > 55 años	56%	2,33 (3,13)		0,83	0,92	100%	84%	85%	>4
Beck et al. (2001)												
Steer												
Ball, Ciervo y Kabat (1997)	56	EEUU	Ambulatorios						83%	95%		>6
Healey et al. (2007)	49	Reino Unido	Personas mayores supervivientes de Infarto	57%	Mediana = 4		Test-retest = 0,63	0,75	71%	74%		

Tabla 4.6. Resumen de las propiedades psicométricas del BDI-FS en la literatura científica (continuación).

Leng, Changrani, Tseng y Gany (2010) realizaron un estudio comparativo sobre la detección de la depresión en 782 pacientes de AP con diferentes lenguas maternas y diferentes usos de interpretación. Encontraron que el BDI-FS en su versión inglesa detectaba depresión del mismo modo en los tres grupos: 1) lengua inglesa, 2) español (origen latino) o 3) chino mandarino o cantonés, los dos últimos con traducción simultánea respectivamente. Como limitaciones del estudio, cabe resaltar que usan como criterio o “patrón oro” el criterio médico y que no aportan datos sobre las propiedades psicométricas del BDI-FS en su muestra.

En español, el trabajo realizado por Lam et al. (2009) en 307 mujeres peruanas gestantes encuentra bajos niveles de consistencia interna, $\alpha = 0,68$, y una estructura bifactorial del BDI-FS, datos ambos que discrepan con los estudios previos. Los autores justifican estos datos tan pobres a la ausencia en el BDI-FS de un componente ansioso “presente en casi todas las depresiones”, pero las deficiencias metodológicas del estudio también podrían explicar esos resultados. Por ejemplo, los investigadores no detallan la versión española que utilizaron ni si dicha versión fue creada específicamente por ellos, por lo que no es posible valorar si la traducción empleada era adecuada para la población objeto de estudio o si las participantes pudieron tener algún problema para comprender el significado de los ítems.

Para poder comparar las características del BDI-FS respecto a las otras dos versiones más utilizadas del BDI, el BDI-IA (1978) y el BDI-II (1996), en la Tabla 4.7 se resumen las principales características de todos ellos.

Tabla 4.7. Comparación de características del BDI-IA, BDI-II y BDI-FS

Características	BDI-IA	BDI-II	BDI-FS
Referencia	Beck et al. (1978)	Beck, Steer y Brown (1996)	Beck, Steer y Brown (2000).
Formato	Autoaplicado	Autoaplicado	Autoaplicado
Nº de ítems	21	21	7
Escala	Alternativas por orden de gravedad (0-3)	Alternativas por orden de gravedad (0-3)	Alternativas por orden de gravedad (0-3)
Marco temporal	Última semana	Dos últimas semanas	Dos últimas semanas
Duración	5-10 minutos	10 minutos	5 minutos
Baremos españoles	Adultos de la población general; pacientes con trastornos psicológicos; estudiantes universitarios	Adultos de la población general; pacientes con trastornos psicológicos; estudiantes universitarios	Sin baremos en español
Fiabilidad	$\alpha = 0,83$	$\alpha = 0,91$	$\alpha = 0,85-0,88$
Validez de contenido	Aceptable: cubre 6/9 criterios del episodio depresivo mayor	Buena: 8/9 criterios del episodio depresivo mayor y 7/7 criterios de distimia.	Aceptable: 6/9 criterios del episodio depresivo mayor y 6/7 criterios de distimia.
Validez convergente	Satisfactoria con HDRS, SDS y ZUNG	Satisfactoria con HDRS, SDS y ZUNG	Satisfactoria con HAD
Validez discriminante	Correlaciones moderadas con BAI o STAI. Análisis factorial de BDI con instrumentos de ansiedad arroja dos factores claros: ansiedad y depresión.	Correlación con BAI = 0,58. Identifica factor depresivo frente a ansioso (BAI)	Correlación con BAI-PC = 0,58
Validez de criterio	Aceptable: discrimina depresivos con no depresivos	Excelente capacidad para diferenciar entre pacientes con diferentes niveles de gravedad y distintos trastornos psicopatológicos	Muy buena capacidad de discriminación entre pacientes depresivos y no depresivos: área bajo curva ROC = 0,94; con punto de corte > 4, sensibilidad = 91% y especificidad = 79%
Validez factorial	Factor I: Cognitivo- afectiva Factor II: Somático-de rendimiento	Factor I: Cognitivo-afectivo Factor II: Somático Factor general de depresión	Factor cognitivo-afectivo
Adaptación española	Vázquez y Sanz (1991, 1997, 1999); Sanz y Vázquez (1998)	Sanz, Navarro y Vázquez (2003); Sanz, Perdigón y Vázquez (2003); Sanz et al. (2005)	En proceso

En la misma línea, en la Tabla 4.8 se presenta la correspondencia de los ítems de las versiones del BDI-IA (1978), del BDI-II (1996) y del BDI-FS (2000) con los criterios diagnósticos sintomáticos del DSM-IV para los trastornos depresivos. En dicha tabla se puede apreciar que las tres versiones cubren de manera aceptable los criterios DSM-IV para los trastornos depresivos y, por tanto, muestran una aceptable validez de contenido, aunque, por supuesto, también hay diferencias entre las versiones respecto a los criterios cubiertos.

En la presente tesis doctoral, se va examinar las propiedades psicométricas de una versión española del BDI-FS creada tomando de la adaptación española anteriormente mencionada del BDI-II, los ítems que componen el BDI-FS, todos los cuales se refieren a síntomas depresivos cognitivos y afectivos. De esta forma, cabe tener en cuenta que la traducción española del BDI-FS procede de un cuidadoso proceso de traducción y adaptación del BDI-II. Ese trabajo se puede consultar en Sanz et al. (2003). De manera resumida, Sanz, Navarro y Vázquez (2003) realizaron un proceso de traducción del BDI-II en varias fases, con varios traductores y contando con el procedimiento de traducción inversa. Finalmente, para asegurarse de la adecuada comprensibilidad de las instrucciones y de sus ítems, Sanz, Navarro y Vázquez (2003) realizaron un estudio piloto del instrumento para evaluar precisamente esos aspectos con una muestra de adultos de la población general sin estudios universitarios.

Tabla 4.8. Correspondencia entre los criterios diagnósticos sintomáticos del DSM-IV para los trastornos depresivos y los ítems del BDI-IA, BDI-II y BDI-FS

<i>DSM IV (1994)</i>	<i>BDI-IA (Beck et al., 1979)</i>	<i>BDI-II (Beck, Steer y Brown, 1996)</i>	<i>BDI-FS (Beck, Steer y Brown, 2000)</i>
1. Estado de ánimo deprimido (o irritable en niños- adolescentes)	(1) Tristeza (10) Lloro (11) Irritabilidad	(1) Tristeza (10) Llanto (17) Irritabilidad	(1) Tristeza
2. Disminución del placer o del interés	(4) Insatisfacción general (12) Desinterés social (21) Desinterés por el sexo	(4) Pérdida de placer (12) Pérdida de interés (21) Pérdida de interés en el sexo	(4) Anhedonia
3. Aumento/ disminución del peso o apetito	(18) Disminución del apetito (19) Disminución del peso	(18) Cambios en el apetito	
4. Insomnio o hipersomnia	(16) Insomnio	(16) Cambios en el patrón de sueño	
5. Agitación o enlentecimiento psicomotor		(11) Agitación	
6. Fatiga o pérdida de energía	(17) Fatiga	(15) Pérdida de energía (20) Cansancio o fatiga	
7. Sentimientos excesivos de inutilidad o culpa	(3) Sentimientos de fracaso (6) Sentimientos de castigo (5) Culpabilidad (8) Autoculpación	(3) Sentimientos de fracaso (6) Sentimientos de castigo (5) Sentimientos de culpa (8) Autocríticas (14) Inutilidad	(3) Sentimiento de fracaso (6) Culpabilidad
8. Disminución de la capacidad para pensar, concentrarse o tomar decisiones.	(13) Indecisión	(13) Indecisión (19) Dificultad de concentración	
9. Ideas recurrentes de muerte o suicidio.	(9) Pensamientos de suicidio	(9) Pensamientos de suicidio	(7) Pensamientos de suicidio
10. Baja autoestima.	(7) Auto decepción (14) Autoimagen	(7) Insatisfacción con uno mismo	(5) Pérdida de autoconfianza
11. Sentimientos de desesperanza	(2) Pesimismo	(2) Pesimismo	(2) Desánimo futuro
Síntomas no recogidos en DSM-IV	(20) Hipocondría		

PARTE EMPÍRICA:

Capítulo 5:

Objetivos e hipótesis de la investigación empírica.

*“La ignorancia afirma o niega
rotundamente; la ciencia duda”*

Voltaire

En los capítulos anteriores de esta tesis doctoral se ha señalado la relevancia sociosanitaria de los trastornos depresivos en AP tanto por su elevada prevalencia como por sus importantes repercusiones negativas, se ha subrayado la necesidad de mejorar su detección por parte de los médicos de AP puesto que las cifras de infradiagnóstico son manifiestamente mejorables, y se ha constatado como la utilización de instrumentos de cribado puede mejorar dicha detección y, por tanto, el cuidado sanitario de los pacientes de AP con depresión. En los capítulos previos también se han revisado los instrumentos de evaluación de la depresión que se suelen utilizar en España y que pueden servir como instrumentos de cribado para los médicos españoles de AP, y se ha constatado que actualmente no se cuenta con ninguna adaptación española del BDI-FS, una versión del BDI-II específicamente diseñada como instrumento para la evaluación de la sintomatología depresiva y el cribado de la depresión en pacientes médicos. Como se ha argumentado en el Capítulo 4, el BDI-II es sin duda el instrumento de autoinforme más utilizado y validado a nivel internacional para la evaluación de la sintomatología depresiva en todo tipo de poblaciones clínicas y no clínicas, y su versión para pacientes médicos, el BDI-FS, presenta una serie de características justamente por proceder del BDI-II que podrían convertirlo en un instrumento tremendamente útil para la evaluación y cribado de la depresión en pacientes españoles de AP, posibilidad que vendría sustentada en las excelentes propiedades psicométricas que justamente para esos objetivos ha demostrado la versión original del BDI-FS en muestras de pacientes de AP de diferentes países de lengua inglesa.

En este contexto, el objetivo general de la investigación empírica de esta tesis doctoral era validar en la población española de pacientes de AP una versión en español del BDI-FS y comprobar específicamente sus propiedades como instrumento de cribado de la depresión en dicha población. Para ello, se llevaron a cabo dos estudios instrumentales con muestras de pacientes de AP de dos ciudades españolas, estudios que se describirán con detalle en los próximos dos capítulos y cuyos objetivos e hipótesis se describen a continuación.

5.1. Objetivos e hipótesis del primer estudio

El objetivo general del primer estudio era analizar las propiedades psicométricas básicas del BDI-FS en una muestra española de pacientes de AP y, en particular, analizar su rendimiento diagnóstico como instrumento de cribado de la depresión y comparar dicho rendimiento diagnóstico con el del juicio clínico de los médicos de AP y el de otro instrumento de cribado de la depresión que está validado y es ampliamente utilizado en España, la subescala de depresión de la HAD.

En el caso de que los resultados que se obtengan sobre las propiedades psicométricas del BDI-FS fueran adecuados, esto permitiría que el BDI-FS pudiera convertirse en un instrumento que ayude de manera importante a los clínicos españoles en la evaluación y detección de los trastornos depresivos en AP.

Las propiedades básicas que se pretendían analizar en este primer estudio fueron las siguientes:

1. La validez factorial.
2. La fiabilidad de consistencia interna.
3. La validez convergente respecto a dos instrumentos validados para la evaluación de la depresión, el módulo de trastornos del estado de ánimo de la entrevista diagnóstica estructurada SCID-I y la subescala de depresión de la HAD.
4. La validez discriminante respecto a un instrumento validado para la evaluación de la ansiedad en pacientes médicos: la subescala de ansiedad de la HAD.
5. La validez de criterio en cuanto a rendimiento diagnóstico respecto al criterio de referencia del diagnóstico de trastorno depresivo realizado por un clínico especialista con experiencia, sobre la base de una entrevista diagnóstica estructurada, la SCID-I.

En relación con esta última propiedad psicométrica, el estudio primero también tenía como objetivo específico comprobar si el BDI-FS era un mejor instrumento de cribado de los trastornos depresivos que la subescala de la HAD y mejoraba la capacidad para detectar trastornos depresivos de los médicos de AP. En concreto, ese objetivo específico era:

6. Comprobar si los índices de validez de criterio o de rendimiento diagnóstico del BDI-FS respecto al criterio de referencia del diagnóstico de trastorno depresivo realizado por un clínico especialista con experiencia sobre la base de una entrevista diagnóstica estructurada (SCID-I) eran superiores para el BDI-FS en comparación a la subescala de depresión de la HAD y al juicio clínico del médico de AP.

Como se ha comentado en un capítulo anterior, la HAD fue diseñada por Zigmond y Snaith (1983) con la finalidad de proporcionar un instrumento para detectar estados de depresión y ansiedad en pacientes de servicios hospitalarios no psiquiátricos. Al igual que el BDI-FS, la HAD cuenta con 7 ítems para evaluar la sintomatología depresiva y cuenta también con una abundante literatura científica que avala sus adecuadas propiedades psicométricas respecto a su fiabilidad de consistencia interna, validez convergente, validez discriminante, validez factorial y validez de criterio respecto al diagnóstico clínico (véase la revisión de Bjelland et al., 2002). En España se cuenta con varias adaptaciones de la HAD (véase Caro e Ibáñez, 1992; Herrero et al., 2003), y sus propiedades psicométricas en pacientes hospitalarios y pacientes de AP son también apropiadas respecto a su fiabilidad de consistencia interna, validez convergente, validez factorial y validez de criterio respecto al diagnóstico clínico (Caro e Ibáñez, 1992; Herrero et al., 2003; Quintana et al., 2003). Resultaba conveniente, pues, evaluar si la versión española del BDI-FS, que cuenta con el mismo número de ítems que la subescala de depresión de la HAD, presentaba un mejor rendimiento diagnóstico a efectos de cribado de la depresión que la subescala de depresión de la HAD, de manera que se pudiera justificar una preferencia de uso del BDI-FS sobre la HAD cuando el objetivo sea el despistaje de la depresión en pacientes españoles de AP.

En relación con el rendimiento diagnóstico (validez de criterio) del juicio clínico del médico de AP sobre la existencia de un trastorno depresivo respecto al diagnóstico de referencia realizado por un especialista en base a una entrevista diagnóstica estructurada (SCID-I) y la posibilidad de que el BDI-FS presente mejores índices de rendimiento diagnóstico y, por tanto, pueda

mejorar la capacidad del médico de AP de detectar trastornos depresivos entre los pacientes españoles de AP, ya se han documentado en los capítulos anteriores de esta tesis doctoral las dificultades de los médicos de AP para obtener unos índices precisos de detección de la depresión (Gabarrón et al., 2004; OMS, 2008; Mitchell, 2009). Estudios previos que se han realizado en España y que han analizado el grado de detección de la depresión de los médicos de AP y de diferentes instrumentos de cribado (p. ej., Bobes et al., 2003; Garcia-Testal, Sancho, Julve, et al., 1998; Puigdemot, Pérez-Blanco y Figueres, 1997; Soler y Pérez, 1997), han utilizado en su mayoría el juicio clínico de un especialista o del propio médico de AP para analizar la concordancia de los resultados del instrumento con el diagnóstico clínico, pero no han utilizado el juicio clínico de un especialista basado en un instrumento diagnóstico estandarizado. En el presente estudio, el uso de la entrevista diagnóstica estructurada SCID-I permitirá, frente a estudios previos realizados en España, una evaluación más adecuada del rendimiento diagnóstico para la depresión del médico de AP y de los instrumentos de cribado analizados (BDI-FS y HAD) y una mejor comparación de los índices de validez de criterio de los tres procedimientos.

De acuerdo con los objetivos generales y específicos mencionados en los párrafos anteriores, en función de los criterios consensuados sobre los valores adecuados para los distintos índices psicométricos de los instrumentos psicológicos (Bollen y Lennox, 1991; Briggs y Cheek, 1986; Cicchetti, 1994; Nunnally y Bernstein, 1995; Prieto y Muñiz, 2000; Streiner y Cairney, 2007; Swets, 1988) y sobre la base de los resultados de los estudios previos sobre las propiedades psicométricas del BDI-FS en muestras de lengua inglesa, en el Estudio 1º se propusieron las siguientes **hipótesis**:

- 1) Se esperaba encontrar que la versión española del BDI-FS en una muestra de pacientes españoles de AP obtuviera una solución unifactorial similar a la encontrada en la literatura científica y que justificara la utilización de la puntuación global del instrumento.
- 2) Se esperaba encontrar que la versión española del BDI-FS en la muestra de pacientes españoles de AP presentara índices buenos y similares a los encontrados en la literatura científica en relación con la fiabilidad de consistencia interna: coeficiente *alfa* de Cronbach $\geq 0,80$.
- 3) Se esperaba encontrar que la versión española del BDI-FS en la muestra de pacientes españoles de AP presentara índices buenos y similares a los encontrados en la literatura científica en relación con la validez convergente: coeficientes de correlación con la SCID-I y la subescala de depresión de la HAD $\geq 0,50$.
- 4) Se esperaba encontrar que la versión española del BDI-FS en la muestra de pacientes españoles de AP presentara índices adecuados y similares a los encontrados en la literatura científica en relación con la validez discriminante respecto a una medida de ansiedad: coeficiente de correlación con la subescala de ansiedad de la HAD $<$ a los coeficientes de correlación con la SCID-I y la subescala de depresión de la HAD, y patrón de saturaciones factoriales de los ítems del BDI-FS significativas ($\geq 0,40$) en un factor de depresión e irrelevantes ($< 0,25$) en un factor de ansiedad tras el análisis factorial conjunto del BDI-FS y de las subescalas de la HAD.
- 5) Se esperaba encontrar que la versión española del BDI-FS en la muestra de pacientes españoles de AP presentara índices buenos y

similares a los encontrados en la literatura científica en relación con el rendimiento diagnóstico para detectar el trastorno depresivo mayor (o validez de criterio respecto al diagnóstico de referencia de trastorno depresivo mayor): área bajo la curva ROC $\geq 0,90$, índices de sensibilidad y especificidad $\geq 0,80$, e índice kappa de concordancia con el criterio $\geq 0,60$.

- 6) Se esperaba encontrar que el rendimiento diagnóstico para detectar el trastorno depresivo mayor (validez de criterio respecto al diagnóstico de referencia de trastorno depresivo mayor) fuera superior para el BDI-FS en comparación a la subescala de depresión de la HAD o el juicio clínico del médico de AP: área bajo la curva ROC del BDI-FS $>$ área bajo la curva ROC de la subescala de depresión de la HAD o del juicio clínico del médico de AP.

5.2. Objetivos e hipótesis del segundo estudio

Dado que la replicación de resultados es una característica fundamental de la ciencia, dado que los resultados del primer estudio podrían tener problemas de validez externa en su generalización a la población de pacientes españoles de AP puesto que se había realizado en una sola ciudad española y dado que en el Estudio 1º no se habían examinado todas las propiedades psicométricas que sería deseable analizar para validar sólidamente el BDI-FS, se decidió diseñar un segundo estudio cuyo objetivo general fue profundizar en el análisis de las propiedades psicométricas del BDI-FS en una nueva muestra española de pacientes de AP.

Más concretamente, los objetivos del estudio segundo eran:

1. Analizar una propiedad psicométrica básica que no fue examinada en el Estudio 1º, la estabilidad temporal o fiabilidad test-retest, y que tampoco ha sido examinada previamente en una muestra de pacientes de AP en ningún estudio publicado nacional o extranjero que se conozca.
2. Ampliar los datos sobre algunas propiedades psicométricas que ya fueron analizadas en el Estudio 1º, en particular, la validez convergente y la validez discriminante, pero extendiéndolas a otros instrumentos que no fueron empleados en el Estudio 1º y que también están validados en España y se usan con frecuencia en AP: la Escala Autoaplicada para la Depresión de Zung (SDS) y el Inventario de Ansiedad de Beck (*Beck Anxiety Inventory* o BAI, Beck, Epstein, Brown y Steer, 1988).

Como se comentó en un capítulo anterior, la SDS fue construida con el objetivo de obtener una medida tanto de la existencia de síntomas depresivos como de su gravedad, y es un instrumento cuya adaptación española cuenta con numerosos datos sobre su fiabilidad y depresión para ese objetivo (véase en el Capítulo 7 la descripción de instrumentos utilizados en el Estudio 2º). Por su parte, el BAI se creó con el objetivo de valorar la presencia de síntomas de ansiedad y su gravedad, pero también con el fin de discriminar de manera más fiable entre ansiedad y depresión, objetivos para los cuales cuenta con índices de fiabilidad y validez muy buenos tanto en su versión original como en su versión española (véase en el Capítulo 7 la descripción de instrumentos utilizados en el Estudio 2º).

3. Replicar en una nueva muestra de pacientes de AP procedentes de otra ciudad española distinta a la ciudad en la que se hizo el Estudio 1º los resultados encontrados en este estudio primero respecto a las

propiedades psicométricas de fiabilidad de consistencia interna y de validez factorial del BDI-FS.

4. Analizar, tras combinar los datos que se obtuvieran en el Estudio 2º con los datos obtenidos en el Estudio 1º, las diferencias en las puntuaciones del BDI-FS en función de las características demográficas básicas (sexo, edad, estado civil y nivel de estudios) de los pacientes de AP y confeccionar unos baremos provisionales del BDI-FS para la población española de pacientes de AP que tuvieran en cuenta esas diferencias y ayudaran a los clínicos a interpretar el nivel de sintomatología depresiva de los pacientes de AP.

De acuerdo con estos objetivos específicos, en función de los criterios consensuados sobre los valores adecuados para los distintos índices psicométricos de los instrumentos psicológicos (Bollen y Lennox, 1991; Briggs y Cheek, 1986; Cicchetti, 1994; Nunnally y Bernstein, 1995; Prieto y Muñiz, 2000; Streiner y Cairney, 2007; Swets, 1988) y sobre la base tanto de los resultados de los estudios previos sobre las propiedades psicométricas del BDI-FS en muestras de lengua inglesa como de los resultados obtenidos en el Estudio 1º, se propusieron en el Estudio 2º las siguientes **hipótesis**:

- 1) Se esperaba encontrar que la versión española del BDI-FS en la nueva muestra de pacientes españoles de AP presentara un índice bueno de fiabilidad de estabilidad temporal a la semana: coeficiente de correlación test-retest $\geq 0,75$.
- 2) Se esperaba encontrar que la versión española del BDI-FS en la nueva muestra de pacientes españoles de AP presentara índices buenos y similares a los encontrados en la literatura científica y en el Estudio 1º en

relación con la validez convergente: coeficientes de correlación con la SDS $\geq 0,50$.

- 3) Se esperaba encontrar que la versión española del BDI-FS en la nueva muestra de pacientes españoles de AP presentara índices adecuados y similares a los encontrados en la literatura científica y en el Estudio 1º en relación con la validez discriminante respecto a una medida de ansiedad: coeficiente de correlación con el BAI $<$ al coeficiente de correlación con la SDS, y patrón de saturaciones factoriales de los ítems del BDI-FS significativas ($\geq 0,40$) en un factor de depresión e irrelevantes ($< 0,25$) en un factor de ansiedad tras el análisis factorial conjunto del BDI-FS, la SDS y el BAI.
- 4) Se esperaba encontrar que la versión española del BDI-FS en la nueva muestra de pacientes españoles de AP presentara índices buenos y similares a los encontrados en la literatura científica y en el Estudio 1º en relación con la fiabilidad de consistencia interna: coeficiente *alfa* de Cronbach $\geq 0,80$.
- 5) Se esperaba encontrar que la versión española del BDI-FS en una nueva muestra de pacientes españoles de AP obtuviera una solución unifactorial similar a la encontrada en la literatura científica y en el Estudio 1º.
- 6) Se esperaba encontrar diferencias en las puntuaciones de la versión española del BDI-FS en muestras de pacientes de AP en función de las características demográficas de dichos pacientes, en concreto, que las mujeres puntuaran significativamente más alto en el BDI-FS que los varones, y que los pacientes viudos, separados o divorciados puntuaran

significativamente más alto en el BDI-FS que los pacientes casados o que convivían con parejas estables, que los pacientes de más edad puntuaran significativamente más alto en el BDI-FS que los más jóvenes, y que los pacientes con menos nivel de estudios puntuaran significativamente más alto en el BDI-FS que los pacientes con mayor nivel de estudios.

Capítulo 6

Primer Estudio: Análisis de las Propiedades Psicométricas del BDI-FS en Atención Primaria en España

6.1. Introducción

Como se ha comentado en el capítulo anterior, el objetivo general de este Estudio 1º fue analizar las propiedades psicométricas del BDI-FS en una muestra de pacientes españoles de AP.

En primer lugar, se quería examinar la validez factorial del BDI-FS, en particular, comprobar si, tal y como se presupone, el instrumento tiene una solución unifactorial en pacientes de AP que avale la puntuación total de depresión que ofrece el instrumento. En el caso de que los resultados factoriales justificasen la validez de la puntuación total del BDI-FS, se pretendía, en segundo lugar, analizar la fiabilidad de consistencia interna de dicha puntuación así como su sensibilidad para detectar diferencias individuales en sintomatología depresiva entre los pacientes de AP.

También se pretendía analizar la validez convergente del BDI-FS respecto a un instrumento que ofreciera una medida sólida de sintomatología depresiva en una muestra de pacientes de AP, para lo cual se eligió la subescala de depresión de la Escala Hospitalaria de Ansiedad y Depresión o HAD (*Hospital Anxiety and Depression Scale*; Zigmond y Snaith, 1993), uno de los instrumentos más utilizados y mejor validados para medir la depresión en pacientes médicos.

De manera complementaria a la validez convergente, en este estudio también se pretendía examinar la validez discriminante del BDI-FS respecto a una medida de un constructo psicopatológico diferente: la ansiedad. Para medir la ansiedad en pacientes de AP también se utilizó la HAD, en particular su subescala de ansiedad, ya que es una de las medidas de ansiedad más

utilizadas y mejor validadas en pacientes médicos (véase la revisión de Bjelland, Dahl, Haug y Neckelmann, 2002).

Finalmente, el Estudio 1º tenía como objetivo específico comprobar la validez de criterio del BDI-FS respecto al diagnóstico de trastorno depresivo ofrecido por un clínico con experiencia mediante la administración de una entrevista diagnóstica estructurada, actualmente el patrón oro de referencia para evaluar el rendimiento diagnóstico de un instrumento. Este objetivo es, sin duda, uno de los más importantes de esta investigación porque tiene que ver de manera más específica con la posibilidad de utilizar el BDI-FS como instrumento de despistaje o cribado (*screening*) en AP, tal y como se ha comentado con detalle en los capítulos anteriores de esta tesis doctoral.

6.2. Método

6.2.1. Participantes

En el presente estudio participó una muestra incidental de pacientes españoles que acuden a AP compuesta por 300 personas, tamaño muestral considerado como adecuado para la validación de instrumentos de evaluación psicológica según los estándares propuestos por Prieto y Muñiz (2000). Sin embargo, algunos participantes dejaron en blanco algunos ítems de los cuestionarios administrados, por lo que la muestra total de la que finalmente se tenían datos osciló entre 295 y 299 participantes. Además, la entrevista diagnóstica estructurada que se describirá más adelante sólo fue administrada a una submuestra de 100 pacientes de esa muestra total. Para cada análisis, se detallará el número exacto de participantes que aportaron datos válidos.

El estudio se llevó a cabo en el centro de atención primaria “Reyes Magos” de Alcalá de Henares. Participaron 9 médicos de atención primaria de turno vespertino. Como el acceso a los pacientes de AP se supone un fenómeno aleatorio, se asignó una hora concreta a cada médico de AP para derivar a un paciente. En el caso que no accediese a participar o no cumpliese los criterios de inclusión, se escogía al paciente siguiente de la lista diaria de citas. A cada médico de AP se le asignó el mismo número de pacientes a derivar.

Se consideraron los siguientes criterios de inclusión: 1) estar adscrito al centro de atención primaria; 2) edad comprendida entre los 18 y los 80 años, y 3) no haber acudido a consulta por motivos exclusivamente administrativos ni por delegación de terceras personas.

Los criterios de exclusión que se utilizaron fueron los siguientes (en función del criterio clínico del médico de AP): 1) tener patología psicótica, demencia o enfermedad concurrente que hiciese imposible la aplicación de las pruebas de estudio, 2) no tener capacidades cognitivas básicas de atención o concentración, y 3) no tener una comprensión suficiente del castellano.

Las principales características sociodemográficas de la muestra se recogen en la Tabla 6.1. La muestra estaba constituida mayoritariamente por mujeres (65%) y la franja etaria más frecuente fue de los 30-44 años en un rango de 18 a 75 años. La mayoría de los participantes estaban casados o convivían de forma estable con una pareja (62%), tenían formación académica de estudios medios (38%) o primarios (35%), estaban trabajando en el momento de realización del estudio (66%). Como se puede apreciar en la Tabla 6.1, las personas que componían la muestra de este estudio tenían profesiones

muy diversas, y sólo tres categorías incluían más de un 10% de la muestra: profesional de las Fuerzas Armadas (13,3%), vendedores y comerciantes (12%) y profesionales técnicos (10,7%).

Tabla 6.1. Características sociodemográficas de la muestra de pacientes de atención primaria del Estudio 1º

<i>Característica</i>	<i>Frecuencia</i>	<i>Porcentaje (%)</i>
Sexo:		
– Mujeres	197	65,7
– Varones	203	34,3
Edad:		
– 18-29 años	80	26,7
– 30-44 años	119	39,7
– 45-59 años	52	17,3
– 60 años o más	28	9,3
– No contesta	21	7,0
Estado civil:		
– Soltero	86	28,7
– Casado o conviviendo con pareja de forma estable	186	62,0
– Separado o divorciado	17	5,7
– Viudo	10	3,3
– No contesta	1	0,3
Formación académica:		
– No tiene estudios	11	3,7
– Estudios primarios, EGB	105	35,0
– Bachillerato, Formación Profesional	114	38,0
– Estudios universitarios	69	23,0
– No contesta	1	0,3

<i>Característica</i>	<i>Frecuencia</i>	<i>Porcentaje (%)</i>
Situación laboral:		
– Trabaja	199	66,3
– Jubilado o pensionista	25	8,3
– Parado	20	6,7
– Estudiante	15	5,0
– Sus labores	33	11,0
– Incapacidad laboral transitoria	3	1,0
– Otra situación	4	1,3
– No contesta	1	0,3
Profesión:		
– Profesionales, técnicos	32	10,7
– Altos directivos de la administración pública/empresa privada	9	3,0
– Propietarios o gerentes de hostelería, comercio, agricultura	8	2,7
– Mandos intermedios, profesores, técnico de la administración	25	8,3
– Capataces y personal encargado en empresas o actividades de servicio	12	4,0
– Personal administrativo	21	6,9
– Comerciantes, vendedores y similares	36	12,0
– Personal de servicio	26	8,7
– Trabajadores cualificados o semicualificados	29	9,7
– Fuerzas Armadas	40	13,3
– Estudiante	1	0,3
– Sus labores	17	5,7
– Otra ocupación	23	7,7
– No contesta	21	7,0

Tabla 6.1. Características sociodemográficas de la muestra de pacientes de atención primaria del Estudio 1º (Continuación)

En relación con las características clínicas de la muestra, los motivos de consulta más frecuentes por los que los pacientes habían acudido al centro de AP fueron problemas musculoesqueléticos (7,7%) y respiratorios (5,3%). Aproximadamente la mitad de los pacientes estaban tomando medicación (49,3%), de manera que un 12% de la muestra estaba tomando psicofármacos, bien ansiolíticos (6.4%) o bien antidepresivos (6.4%). Véase la Tabla 6.2 para una descripción más detallada de las características clínicas de la muestra respecto al motivo de consulta y al tipo de medicación.

Tabla 6.2. Características clínicas de la muestra de pacientes de atención primaria del Estudio 1º

Característica clínica	Frecuencia	Porcentaje (%)
Motivo de consulta* al MAP:		
– Infecciosas	8	2,7
– Endocrinas	6	2,0
– Sangre	7	2,3
– Sistema nervioso central	11	3,7
– Cardiovascular	4	1,3
– Respiratorias	16	5,3
– Gastrointestinal	9	3,0
– Musculoesquelética	23	7,7
– Neoplasias	1	0,3
– Ginecología/urología	6	2,0
– Otros	9	3,0
¿Está tomando medicación?:		
– No	151	50,3
– Sí	148	49,3
– No contesta	1	0,3
Tipo de medicación:		
– Ansiolítico	19	6,4
– Antidepresivo	19	6,4
– Antihipertensivo	21	7,1
– Analgésico	12	4,1
– Otros	72	24,4
– Combinación antidepresivo y ansiolítico	6	2,0
– Combinación varios	1	0,3

Nota. * Según la *International Classification of Diseases and Related Health Problems, Tenth Edition, Revised* (ICD-10) de la Organización Mundial de la Salud. MAP = médico de atención primaria.

6.2.2. Instrumento de evaluación

- **Inventario de Depresión de Beck–Cribado Rápido o BDI-FS** (Beck et al., 2000). El instrumento objeto de esta tesis ya ha sido descrito con detalle en el Capítulo 4, tanto en su versión original como en la versión española que se está validando.

- **Entrevista Clínica Estructurada para los Trastornos del Eje I del DSM-IV o SCID-I** (First, Spitzer, Gibbon y Williams, 1997). Se trata de una entrevista semiestructurada destinada a establecer los diagnósticos más importantes del Eje I del DSM-IV (*American Psychiatric Association* o APA, 1994). Las entrevistas clínicas estructuradas se crearon para mejorar la fiabilidad diagnóstica mediante la estandarización del proceso de evaluación, la aplicación de criterios diagnósticos consensuados como, por ejemplo, los del DSM-IV, y la indagación sistemática de síntomas que de otra manera pudiesen pasar desapercibidos. Inicialmente, se creó la SCID para el DSM-III-R (APA, 1987), que fue publicada en 1990. Posteriormente, en 1993, se inició una revisión de la SCID de acuerdo a los criterios DSM-IV, finalizándose la versión para investigación en 1996. En 1997 se realizó la primera publicación de la versión clínica original en lengua inglesa. La edición en español apareció publicada en 1999.

Como se ha referido anteriormente, y con fines operativos, se realizaron tres versiones: 1) versión clínica, 2) versión para la investigación, y 3) versión para los trastornos de la personalidad (SCID-II).

La diferencia principal entre la versión clínica y la versión de investigación es que la segunda es más larga dado que incluye la valoración de

los subtipos y especificaciones en gravedad, curso, etc., de los trastornos, mientras que la versión clínica cubre únicamente los diagnósticos DSM-IV *que se observan más habitualmente en la práctica clínica* (en la mayoría aparecen todos los criterios DSM-IV, sin embargo en algunos aparecen de forma resumida). La versión de los trastornos de personalidad, como su propio nombre indica, comprende la evaluación de todo el eje II del DSM-IV.

La SCID en su versión clínica contiene dos tipos de entrevista o módulos:

1. Visión general: Constituida por una entrevista de exploración del sujeto, su entorno y sus problemas clínicos.
2. Seis entrevistas específicas que constituyen otros tantos módulos, cada uno de ellos destinados a evaluar un tipo de trastorno del eje I del DSM-IV.

En esta investigación se utilizó como prueba diagnóstica de referencia el módulo de trastornos del estado de ánimo (módulo A) de la versión clínica de la SCID-I, en su traducción española oficial (First, Spitzer, Gibbon y Williams, 1999). Por tanto, se obtuvo información de los siguientes diagnósticos que incluye dicho módulo: episodio depresivo mayor actual (F32.x), episodio maniaco actual (F30.x), episodio hipomaniaco actual (F31.0), trastorno distímico actual (F34.1), trastorno del estado de ánimo inducido por sustancias (F10.8), y trastorno del estado de ánimo debido a enfermedad médica (F06.xx).

La SCID-I presenta índices adecuados de fiabilidad entre jueces independientes que varían en los estudios desde un kappa de 0,70 a 1 (First et al., 1996), datos muy aceptables acorde a los estándares para la fiabilidad definida como concordancia entre jueces (Cicchetti, 1994), de manera que

suele producir diagnósticos más fiables, precisos y válidos que las entrevistas clínicas normales. (Fernández-Ballesteros et al., 2010). En el caso concreto del trastorno depresivo mayor, los índices kappa de fiabilidad entre jueces se sitúan entre 0,66 (Lobbestael, Leurgans y Arntz, 2011) y 0,80 (Zanarini et al., 2001), y en el caso del trastorno distímico entre 0,76 (Zanarini et al., 2001) y 0,81 (Lobbestael et al., 2011).

- **Escala Hospitalaria de Ansiedad y Depresión o HAD** (*Hospital Anxiety and Depression Scale*; Zigmond y Snaith, 1983). Esta es una escala autoaplicada diseñada para la detección de trastornos depresivos y ansiosos en el contexto hospitalario no psiquiátrico. La HAD está constituida por 14 ítems, ninguno de los cuales hace referencia a síntomas somáticos. Los ítems se agrupan en dos subescalas, cada una de ellas con 7 ítems:

- Subescala de ansiedad: está centrada en sus manifestaciones psíquicas (tensión, nerviosismo, aprensión, preocupación, inquietud, nervios en el estómago, angustia). Son los ítems impares de la HAD.

- Subescala de depresión: centrada en la anhedonia (disfrute, risa, alegría, torpeza, interés por el aspecto personal, ilusión). Son los ítems pares de la HAD.

El paciente ha de contestar a cada ítem utilizando una escala de tipo Likert que unas veces hace referencia a la intensidad del síntoma y otras veces a la frecuencia de su presentación. En cualquier caso, la escala oscila entre 0 (nunca) y 3 (casi todo el día, muy intenso). El marco de referencia temporal es la semana previa.

En España existen al menos tres traducciones al español de la HAD que han sido validadas en muestras de pacientes españoles con diversos trastornos de salud. Por un lado, se cuenta con la versión traducida por Snaith, Bulbena y Berrios y validada en una muestra de pacientes psiquiátricos por Tejero et al. (1986; véase también De las Cueva Castresana, García-Estrada Pérez y González de Rivera, 1995), en una muestra de pacientes médicos ambulatorios por Herrero, Blanch, Peri, De Pablo, Pintor y Bulbena (2003), en una muestra mixta de pacientes médicos y psiquiátricos, estudiantes universitarios y adultos de la población general por Quintana, Padierna, Esteban, Arostegui, Bilbao y Ruiz (2003), en una muestra de pacientes oncológicos por López-Roig et al. (2000) y en una muestra mixta de estudiantes universitarios y adultos de la población general por Terol, López-Roig, Rodríguez-Marín, Martín-Aragón, Pastor y Reig (2007). También se cuenta con la versión traducida por Caro e Ibáñez (1992) y validada en una muestra de adultos de población general (Caro e Ibáñez, 1992) y en una muestra de pacientes con dolor crónico (Monsalve, Soriano e Ibáñez, 2001). Por último, en Bobes, Portilla, Bascarán, Sáiz y Bousoño (2002) aparece una tercera traducción española de la HAD que ha sido validada en una muestra de pacientes con lesión medular (Villareal Salcedo, 2008) y ha sido recomendada por la “Guía de práctica clínica para el manejo de pacientes con trastornos de ansiedad en atención primaria” del Ministerio de Sanidad y Consumo español (Grupo de Trabajo de la Guía de Práctica Clínica para el Manejo de Pacientes con Trastornos de Ansiedad en Atención Primaria, 2008). Esta última versión española fue la utilizada en el presente estudio.

Todas las versiones españolas de la HAD mencionadas anteriormente, incluyendo la utilizada en el presente estudio, son muy semejantes entre sí y presentan propiedades psicométricas adecuadas en cuanto a su fiabilidad y validez. Por ejemplo, en una muestra de pacientes españoles médicos con diversas patologías la HAD presentó índices de fiabilidad de consistencia interna (α de Cronbach) de 0,84 para la subescala de depresión y de 0,85 para la subescala de ansiedad (Herrero et al., 2003). En esa misma muestra, la subescala de depresión de la HAD mostró índices de especificidad que rondaban el 87% y de sensibilidad alrededor del 72% para detectar trastornos del estado de ánimo respecto al diagnóstico de referencia ofrecido por la SCID-I (con puntuaciones de corte entre 4 y 7 se obtuvieron los índices más altos de especificidad o sensibilidad).

En la muestra de pacientes de AP del presente estudio, ambas subescalas de la HAD obtuvieron coeficientes *alfa* de Cronbach iguales a 0,82, los cuales pueden considerarse buenos índices de consistencia interna según los estándares propuestos por Prieto y Muñiz (2000) y similares a los obtenidos por Herrero et al. (2003) en su muestra española de pacientes médicos con diversas patologías.

6.2.3. Procedimiento

El estudio se realizó en dos fases que se resumen en la Figura 6.1.

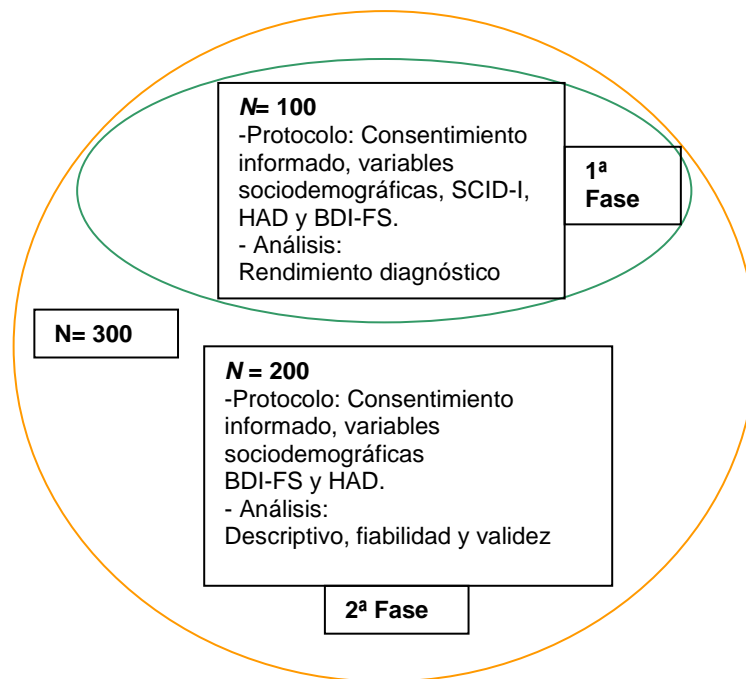
1) Fase 1: En esta fase se evaluaron 100 pacientes que fueron elegidos de forma sistemática en las consultas de AP mediante la selección de las citas que correspondían a las horas en punto y a las citas de las medias horas. A

estos pacientes, el médico de AP le expuso la existencia de una investigación para la exploración de alteraciones emocionales y solicitó su colaboración. En aquellos pacientes que quisieron participar, su médico de AP le dio una hoja cumplimentada con el motivo de consulta actual, el juicio clínico actual y las enfermedades diagnosticadas que estaban en relación con alteraciones psiquiátricas en los últimos 6 meses. La identificación de la persona se hizo con una clave formada por el número de paciente enviado de forma correlativa y el número de identificación asignado al médico de AP que lo derivaba, para que así la investigadora que administraba la entrevista clínica no conociera con anticipación el juicio del médico de AP. Posteriormente el médico de AP derivó al paciente, con la hoja de diagnóstico, al despacho de la investigadora. Ésta le explicó al paciente la naturaleza del estudio y obtuvo el consentimiento informado para su participación, advirtiéndole que podía rechazar seguir con la investigación en cualquier momento de la misma sin ningún tipo de perjuicio. Posteriormente, se facilitó al paciente un protocolo formado por el BDI-FS, la HAD y un cuestionario de variables sociodemográficas, y una vez finalizados estos cuestionarios se le administró la entrevista diagnóstica SCID-I.

2) Fase 2: Se seleccionaron otros 200 pacientes de AP en las mismas consultas que se seleccionaron los primeros 100 pacientes de la muestra total. A estos pacientes, la investigadora les explicó individualmente, durante el tiempo de espera a su cita con su médico de AP, la naturaleza del estudio y obtuvo de ellos el consentimiento informado de participación en la investigación. Posteriormente, se hizo entrega a los pacientes de los cuestionarios (BDI-FS, HAD y cuestionario de variables sociodemográficas),

indicando el modo de cumplimentación, y supervisando y ayudando al participante que lo requería.

Figura 6.1. Resumen del procedimiento del Estudio 1º



La investigadora que llevó a cabo las dos fases del procedimiento del estudio había recibido previamente entrenamiento específico en el uso del BDI-FS, de la HAD y de la SCID-I, este último siguiendo la metodología propuesta en el Manual del Usuario de la entrevista (First et al., 1999). La investigadora y la instrumentación (cuestionarios) fueron iguales para todos los participantes.

6.2.4. Diseño y análisis de datos

Este estudio respondió a un diseño correlacional transversal de tipo instrumental. Para el análisis de los datos obtenidos, se utilizó el paquete

estadístico *SPSS* 15.0 para *Windows XP*. Para analizar de forma descriptiva las distribuciones de las puntuaciones y las variables sociodemográficas se calculó la media, la desviación típica y los índices de curtosis y simetría.

La validez factorial se analizó mediante análisis factorial a través del método de extracción por factorización de ejes principales. La fiabilidad se analizó a través del coeficiente alfa (α) de consistencia interna de Cronbach. La validez convergente y discriminante se analizó mediante el cálculo de los coeficientes de correlación del BDI-FS con las subescalas de la HAD y calculando las diferencias entre dichos coeficientes de correlación mediante la prueba *Z* de Steiger. Así mismo, para analizar la validez discriminante del BDI-FS, se realizó un análisis factorial conjunto de todos los ítems del BDI-FS y de la HAD según el método de extracción de factores de ejes principales y de rotación oblicua de factores *promax*.

Para analizar la validez de criterio del BDIF-FS, en concreto su rendimiento diagnóstico respecto al diagnóstico de trastorno depresivo mayor ofrecido por el instrumento de referencia, la SCID-I, se utilizó el programa estadístico *MedCalc* para *Windows XP*, versión 11.6.0.0 (*MedCalc Software*, Mariakerke, Bélgica). A partir de este programa se calcularon los siguientes índices de rendimiento diagnóstico para las diferentes puntuaciones del BDI-FS, de la subescala de depresión de la HAD y del juicio clínico del médico de AP respecto a la existencia de un trastorno depresivo: sensibilidad (probabilidad de que el despistaje sea positivo cuando efectivamente existe un trastorno depresivo mayor, es decir, porcentaje de pacientes con un trastorno depresivo mayor que puntúan positivo en el cribado), especificidad (probabilidad de que el cribado sea negativo cuando realmente no está presente el trastorno depresivo,

es decir, porcentaje de pacientes sin trastorno depresivo mayor que puntúan negativo en el cribado), valor predictivo positivo (probabilidad de que, ante un caso positivo en el cribado, el paciente sufra realmente de un trastorno depresivo, es decir, la proporción de pacientes con un resultado positivo en la prueba que finalmente padecen el trastorno), valor predictivo negativo (probabilidad de que, ante un caso negativo en el cribado, el paciente no sufra realmente de un trastorno depresivo, es decir, la proporción de pacientes con un resultado negativo en la prueba que finalmente no padecen el trastorno), eficiencia diagnóstica o índice de validez (proporción de pacientes con y sin episodio depresivo mayor clasificados correctamente), índice de Youden (una medida conjunta de eficiencia diagnóstica basada en la sensibilidad y la especificidad que refleja la diferencia entre la tasa de verdaderos positivos y la de falsos positivos), e índice *kappa* de concordancia (grado de acuerdo entre el cribado y el diagnóstico ofrecido por SCID-I que no es explicable por azar).

Con el objetivo de analizar el rendimiento diagnóstico a lo largo de todas las puntuaciones del BDI-FS y de la subescala de depresión de la HAD, se empleó también un análisis de la curva ROC (curva de la característica operativa del receptor) o curva de rendimiento predictivo, realizado también a partir del programa *MedCalc*. En la curva ROC se representa gráficamente la sensibilidad en la ordenada y la tasa de falsos positivos ($1 - \text{especificidad}$) en la abscisa para cada uno de los puntos de corte del instrumento. El caso ideal estaría representado por un punto de corte que arrojase una sensibilidad del 100% y una tasa de falsos positivos de 0% (especificidad del 100%); lo cual quedaría representado en el vértice superior izquierdo de la curva. De este modo, el mejor punto de corte para el instrumento en cuestión sería aquel que estuviese situado

más cerca de dicho vértice. En el caso de que el poder discriminante del instrumento fuese nulo, los valores de la curva ROC formarían una línea recta en la diagonal principal. El programa *MedCalc* también permitió calcular el área bajo la curva como índice resumen del rendimiento diagnóstico para cada uno de los instrumentos o procedimientos de cribado (BDI-FS, subescala de depresión de la HAD y juicio diagnóstico médico) y comparar por pares las diferencias en el área bajo la curva. Para este último análisis se empleó el estadístico de contraste Z de Steiger, también calculado con *MedCalc*.

6.3. Resultados

6.3.1. Prevalencia de los trastornos depresivos

A partir del módulo de trastornos del estado de ánimo de la entrevista estructurada SCID-I se obtuvo en la submuestra a la que se aplicó dicha entrevista ($n = 100$) una estimación de la prevalencia de los trastornos depresivos en los pacientes de AP que participaron en el Estudio 1º. En concreto, con la SCID-I, se encontró un 20% de pacientes con trastorno depresivo mayor (20 pacientes), un 0,3% con trastorno distímico (1 paciente) y un 3% (3 pacientes) con un trastorno del estado de ánimo debido a una enfermedad médica. Los restantes 76 pacientes de la submuestra (76%) no presentaban ningún trastorno del estado de ánimo de acuerdo a la SCID-I.

6.3.2. Validez factorial

Se realizó un análisis factorial sobre las puntuaciones de los siete ítems del BDI-FS. Previamente, y para comprobar la viabilidad del análisis factorial para los datos del presente estudio, se calcularon dos índices: el índice de adecuación muestral de Kaiser-Meyer-Olkin (KMO) y la prueba chi-cuadrado de esfericidad de Bartlett. Ambos índices revelaron que el modelo factorial era adecuado para esos datos. El KMO arrojó un valor de 0,88, valor que superaba el estándar de 0,60 que se considera el mínimo para estimar una buena adecuación de los datos muestrales al análisis factorial. La prueba de esfericidad de Bartlett fue estadísticamente significativa [chi-cuadrado (21) = 724,60, $p < 0,001$], con lo que se rechazó la hipótesis nula de que la matriz de correlaciones entre los ítems del BDI-FS fuera una matriz identidad, es decir, la hipótesis nula de la no existencia de correlaciones entre los ítems del BDI-FS.

Tabla 6.3. Autovalores y porcentajes de varianza de los factores extraídos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 1º

<i>Factores</i>	<i>Autovalor</i>	<i>Porcentaje de la varianza</i>
1	3,729	53,3
2	,734	10,5
3	,652	9,3
4	,574	8,2
5	,523	7,5
6	,417	5,9
7	,370	5,3

Nota. $N = 295$

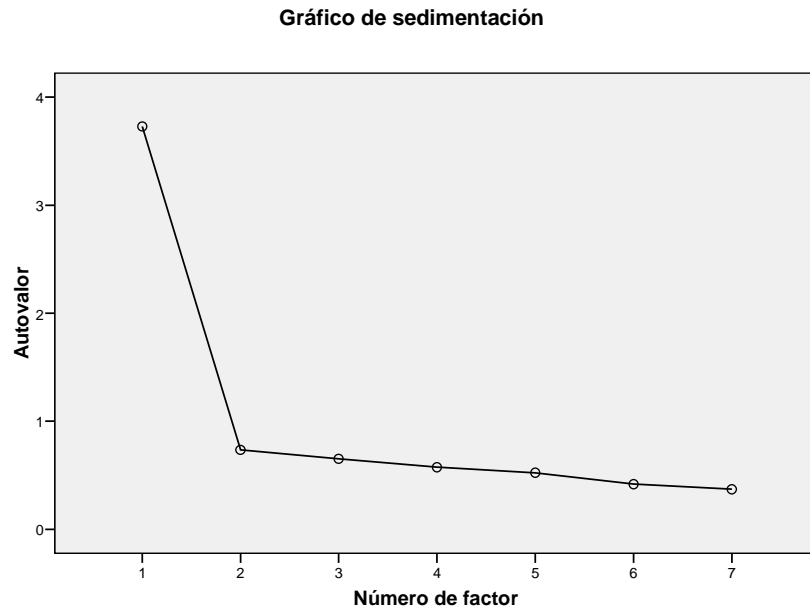
La extracción inicial de factores con factorización de ejes principales arrojó solo un factor con un autovalor o valor propio superior a 1. Este factor explicaba el 53,3% de la varianza, mientras que los siguientes factores explicaban individualmente menos de un 11% de varianza (véase la Tabla 6.3). El gráfico de sedimentación (*scree test*) de Cattell sugería también una solución unifactorial (véase la Figura 6.2), al igual que lo hacía el test de mínima correlación parcial promediada de Velicer (MAP) realizado mediante el programa de instrucciones de SPSS de O'Connor (2000). Los estudios de simulación muestran que el MAP es uno de los mejores métodos para determinar el número de factores (Zwick y Velicer, 1986). Es más, en la matriz factorial de un único factor todos los ítems del BDI-FS saturaban en él más de 0,50 (véase la Tabla 6.4).

Tabla 6.4. Matriz factorial del BDI-SF en la muestra de pacientes del Estudio 1º tras la extracción de un factor mediante el método de ejes principales

<i>Ítem del BDI-FS</i>	<i>Factor 1</i>
Tristeza	,742
Pérdida de placer	,707
Insatisfacción con uno mismo	,684
Sentimientos de fracaso	,683
Pesimismo	,677
Pensamientos o deseos de suicidio	,635
Autocríticas	,589

Nota. N = 295

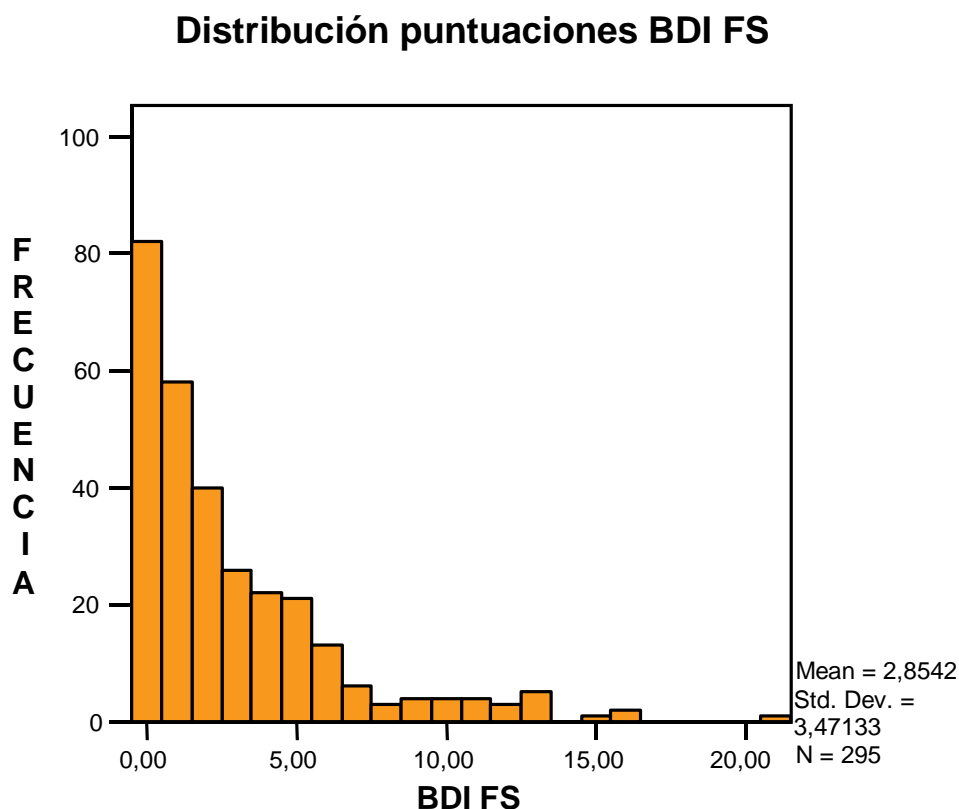
Figura 6.2. Gráfico de sedimentación de Cattell de los factores obtenidos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 1º



6.3.3. Distribución de las puntuaciones del BDI-FS

Las puntuaciones totales en el BDI-FS de la muestra de pacientes de AP del Estudio 1º ($N = 295$) oscilaron entre 0 y 21, con una media igual a 2,85 y una desviación típica de 3,47.

Figura 6.3. Distribución de las puntuaciones del BDI-FS en la muestra de pacientes de atención primaria del Estudio 1º



Diversos indicadores estadísticos y gráficos (véase la Figura 6.3) sugerían que las puntuaciones del BDI-FS en la presente muestra de pacientes de AP no se distribuían de forma normal. El test de normalidad de Kolmogorov-Smirnov fue estadísticamente significativo ($D = 0,207$, $p < 0,001$), lo que implicaba que la distribución del BDI-FS se desviaba de forma estadísticamente significativa de la distribución teórica normal. Además, los índices obtenidos de curtosis y asimetría, con valores de 4,4 y 1,9, respectivamente, se encontraban fuera del rango de los valores que indican una distribución normal (± 1) y señalaban que la distribución de puntuaciones en el BDI-FS se encontraba

desplazada hacia la izquierda (con un mayor número de casos al lado izquierdo de la media) y apuntada en los valores más bajos con una concentración en esos valores mucho mayor de lo que se esperaría de una distribución normal (véase la Figura 6.3).

Tabla 6.5. Distribución de la muestra de pacientes de atención primaria del Estudio 1º en función de los niveles de gravedad del BDI-FS

<i>Rangos cuantitativos de depresión</i>	<i>Frecuencia</i>	<i>Porcentaje</i>
Depresión mínima o ausente (BDI-FS = 0-3)	206	68,7
Depresión leve (BDI-FS = 4-6)	56	18,7
Depresión moderada (BDI-FS = 7-9)	13	4,3
Depresión grave (BDI-FS = 10-21)	20	6,7
Total de pacientes	295	

La distribución de la muestra de pacientes en función de las categorías de gravedad de la depresión propuestas en el manual del BDI-FS (Beck et al., 2000) se presenta en la Tabla 6.5. Como cabría esperar de los índices de distribución de las puntuaciones del BDI-FS mencionados anteriormente, la mayoría de los pacientes (68,7%) obtuvieron puntuaciones bajas en el BDI-FS (entre 0 y 3), en el rango de puntuaciones que Beck et al. (2000) consideran indicadores de depresión mínima o ausente.

En la Tabla 6.6 se presentan para cada ítem del BDI-FS la media, la desviación típica y los porcentajes de respuesta a cada una de sus alternativas de respuesta. La puntuación media de todos los ítems del BDI-FS fue 0,40 con un mínimo de 0,11 y un máximo de 0,73.

Teniendo en cuenta solamente a los pacientes con un diagnóstico de trastorno depresivo mayor según la SCID-I ($n = 20$), su puntuación media en el BDI-FS fue igual a 9 ($DT = 4,25$) y la distribución de sus puntuaciones no se desviaba de forma estadísticamente significativa de la distribución teórica normal (D de Kolmogorov-Smirnov = 0,16, $p = 0,196$).

Tabla 6.6. Características psicométricas de los ítems del BDI-FS en la muestra de pacientes de atención primaria del Estudio 1º ($N = 295$)

Ítem del BDI-FS	Gravedad del síntoma -				Media	DT	r _{ítem-total}
	% de respuesta						
	0	1	2	3			
1. Tristeza	72,6	21,0	4,3	2,0	,34	,64	,66
2. Pesimismo	68,0	23,3	4,0	3,7	,42	,74	,61
3. Sentimientos de fracaso	74,0	20,0	4,3	1,0	,32	,60	,62
4. Pérdida de placer	60,7	29,3	6,7	3,0	,51	,74	,66
5. Insatisfacción con uno mismo	71,3	19,3	4,3	4,0	,41	,75	,62
6. Autocríticas	45,0	42,0	5,0	7,3	,73	,85	,54
7. Pensamientos de suicidio	89,7	7,7	1,0	0,7	,11	,40	,58

6.3.4. Fiabilidad de consistencia interna

El análisis de la consistencia interna del BDI-FS a través del coeficiente alfa de Cronbach arrojó un valor de 0,85, lo que se considera un buen índice de consistencia interna según los estándares propuestos por Prieto y Muñiz (2000).

El alfa de Cronbach es un índice de consistencia interna que depende del número de ítems del instrumento. Para comparar instrumentos con distintos número de ítems en cuanto a su consistencia interna se ha propuesto utilizar la correlación media entre ítems, ya que ésta no depende del número de ítems del instrumento en cuestión. La correlación mínima entre ítems del BDI-FS en la muestra de pacientes de AP de este estudio fue 0,35 y la máxima fue 0,56, siendo la correlación media entre ítems igual a 0,45, valor que se sitúa cerca del límite más alto del rango de 0,20-0,50 que es considerado como el adecuado para la correlación media entre ítems (véase los estándares propuestos por Briggs y Cheek, 1986, y las matizaciones introducidas por Bollen y Lennox, 1991).

La correlación corregida de cada ítem del BDI-FS con el total de la escala, es decir, la correlación del ítem con la puntuación total del BDI-FS calculada sin tener en cuenta el ítem en cuestión, se muestran en la Tabla 6.6. Todas esas correlaciones corregidas, las cuales pueden entenderse como índices de consistencia interna de los ítems, fueron en todos los casos, superiores al estándar de 0,30 (Nunnally y Bernstein, 1995), oscilando entre 0,54 para el ítem de Sentimientos de culpa y 0,66 para el ítem de Tristeza.

6.3.5. Validez convergente

Las correlaciones entre las puntuaciones del BDI-FS, la HAD y la SCID-I, puntuado este último instrumento como presencia o ausencia de episodio depresivo mayor según los resultados de su módulo correspondiente, se muestran en la Tabla 6.7.

Con respecto a la validez convergente, y tal y como puede observarse en la Tabla 6.7, el BDI-FS correlacionó de forma positiva y estadísticamente significativa tanto con la medida de episodio depresivo mayor obtenida a partir del módulo de episodio depresivo mayor de la SCID-I ($r = 0,77$, $p < 0,001$) como con la medida de sintomatología depresiva obtenida a partir de la subescala de depresión de la HAD ($r = 0,74$, $p < 0,001$). Además, en ambos casos las correlaciones fueron superiores al estándar de 0,60 que Prieto y Muñiz (2000) consideran un índice excelente de validez convergente.

Tabla 6.7. Correlaciones entre las medidas de depresión y ansiedad en la muestra de pacientes de atención primaria del Estudio 1º

		<i>Episodio depresivo mayor (SCID-I)</i>	<i>Subescala de depresión de la HAD</i>	<i>BDI-SF</i>
<i>Subescala de depresión de la HAD</i>	<i>r</i>	,55		
	<i>N</i>	100		
<i>BDI-FS</i>	<i>r</i>	,77	,74	
	<i>N</i>	100	291	
<i>Subescala de ansiedad de la HAD</i>	<i>r</i>	,57	,66	,65
	<i>N</i>	99	290	291

Nota. Todas las correlaciones fueron significativas con $p < 0,001$ (bilateral).

6.3.6. Validez discriminante

En relación con la validez discriminante del BDI-FS, éste correlacionó de forma positiva y estadísticamente significativa con la subescala de ansiedad del HAD ($r = 0,65$, $p < 0,001$). Aunque la magnitud de esa correlación puede

considerarse grande en función de los estándares de Cohen (1988), sin embargo, es importante señalar, en relación con la validez discriminante del BDI-FS, que las correlaciones que presentó el BDI-FS con la medida de episodio depresivo mayor de la SCID-I y con la subescala de depresión de la HAD fueron mayores que la correlación que presentó el BDI-FS con la subescala de ansiedad de la HAD (0,77 y 0,74 frente a 0,65) (véase la Tabla 6.7).

De hecho, la diferencia entre la correlación del BDI-FS con la subescala de depresión de la HAD y la correlación del BDI-FS con la subescala de ansiedad de la HAD fue estadísticamente significativa [Z de Steiger ($N = 291$) = 2,64, $p = 0,008$]. Sin embargo, no fue posible determinar con seguridad si la diferencia entre la correlación del BDI-FS con la medida de episodio depresivo mayor de la SCID-I y la correlación del BDI-FS con la subescala de ansiedad de la HAD fue estadísticamente significativa ya que, para realizar la prueba estadística correspondiente, se calcularon las correlaciones para la submuestra de 99 pacientes que habían completados los tres instrumentos y, en esta submuestra, la correlación del BDI-FS con la medida de episodio depresivo mayor de la SCID-I fue igual que en la muestra total ($r = 0,77$), pero la correlación del BDI-FS con la subescala de la HAD fue mayor que en la muestra total ($r = 0,71$ frente a 0,65) y, por tanto, la diferencia entre ambas correlaciones del BDI-FS fue menor en la submuestra ($r = 0,77$ frente a 0,71) y no estadísticamente significativa [Z de Steiger ($N = 99$) = 0,96, $p = 0,333$].

Al comparar los índices de validez discriminante del BDI-FS anteriormente mencionados con los obtenidos por la subescala de depresión de la HAD, estos últimos parecieron inferiores. Como puede observarse en la

Tabla 6.7, la correlación de la subescala de depresión de la HAD con la subescala de ansiedad de la HAD fue de 0,66, pero, al contrario de lo que ocurría con el BDI-FS, la correlación de la subescala de depresión de la HAD con la medida de episodio depresivo mayor de la SCID-I no sólo no fue superior a su correlación con la subescala de ansiedad de la HAD, sino que fue más baja (0,55 frente a 0,66).

A pesar de que los datos de la Tabla 6.7 sugieren una aceptable validez convergente-discriminante del BDI-FS, es cierto que la existencia de una correlación discriminante respecto a la subescala de ansiedad de la HAD de una magnitud superior a 0,50, valor que se considera un efecto de tamaño grande según los estándares de Cohen (1988), podría, en principio, sugerir falta de validez discriminante por parte del BDI-FS, salvo que se pudiera demostrar que, a pesar de esa alta correlación, el BDI-FS puede medir una sintomatología afectiva diferente de la ansiedad.

Para responder a esta última cuestión se realizó un análisis factorial con todos los ítems del BDI-FS y de las subescalas de la HAD, ya que además tanto el test de Kaiser-Meyer-Olkin de adecuación muestral ($KMO = 0,93$) como la prueba de esfericidad de Bartlett ($\chi^2(210) = 2638,88, p < 0,001$) indicaban que el modelo factorial era adecuado para esos datos. Puesto que se trataba de examinar si el BDI-FS medía síntomas afectivos diferentes de la ansiedad, en el análisis factorial se extrajeron dos factores mediante ejes principales y se rotaron con el método oblicuo *promax*. Como puede verse en la Tabla 6.8, la matriz de configuración resultante reflejó claramente la existencia de un factor de depresión y otro de ansiedad, los cuales correlacionaban entre

sí de forma alta ($r = 0,74$), pero quedaban claramente definidos por los ítems de sus correspondientes instrumentos.

Tabla 6.8. Análisis factorial conjunto de los ítems del BDI-FS y de la HAD: solución bifactorial.

Ítem (numeración del instrumento)	Instrumento	Factor 1	Factor 2
1. Tristeza	BDI-FS	0,735	–
2. Pesimismo	BDI-FS	0,655	–
3. Sentimientos de fracaso	BDI-FS	0,615	–
4. Pérdida de placer	BDI-FS	0,672	–
7. Insatisfacción con uno mismo	BDI-FS	0,659	–
8. Auto-críticas	BDI-FS	0,426	–
9. Pensamientos o deseos de suicidio	BDI-FS	0,780	–
2. Capacidad de disfrutar	HAD-D	0,670	–
4. Capacidad para reírse	HAD-D	0,495	–
6. Alegría	HAD-D	0,515	0,255
8. Lentitud psicomotora	HAD-D	0,656	–
10. Pérdida de interés por aspecto físico	HAD-D	0,492	–
12. Optimismo	HAD-D	0,466	–
14. Capacidad para divertirse con actividades de ocio	HAD-D	–	0,271
1. Tensión o nerviosismo	HAD-A	–	0,778
3. Miedo a que suceda algo horrible	HAD-A	0,335	0,388
5. Preocupaciones	HAD-A	–	0,645
7. Tranquilidad y relajación	HAD-A	–	0,604
9. Aleteo en el estómago	HAD-A	–	0,555
11. Inquietud psicomotora	HAD-A	–	0,700
13. Pánico	HAD-A	–	0,558

Nota. $N = 288$. HAD-D = subescala de depresión de la Escala Hospitalaria de Ansiedad y Depresión (HAD); HAD-A = subescala de ansiedad de la Escala Hospitalaria de Ansiedad y Depresión (HAD). No se muestran las saturaciones $< 0,25$ y las saturaciones $> 0,40$ se muestran en negrita.

Efectivamente, considerando como despreciables las saturaciones menores de 0,25 en la matriz de configuración y como definitorias las saturaciones iguales o mayores de 0,40 en dicha matriz, todos los ítems que

definían el primer factor (depresión) pertenecían al BDI-FS, cuyos ítems saturaban entre 0,43 y 0,78, o a la subescala de depresión de la HAD, cuyos ítems saturaban entre 0,46 y 0,67, salvo por el ítem 14 (Capacidad para disfrutar con un buen libro, programa de radio o televisión) que no alcanzó a saturar más de 0,25 en el factor de depresión. Es más, todos los ítems del BDI-FS presentaban saturaciones mayores de 0,40 en ese primer factor de depresión (véase la Tabla 6.8).

De forma similar, todos los ítems que definían el segundo factor (ansiedad) pertenecían a la subescala de ansiedad de la HAD, cuyos ítems saturaban entre 0,55 y 0,78, salvo por el ítem 3 (Miedo de que suceda algo horrible) que, casi por una décima no alcanzó a saturar 0,40 o más en el factor de ansiedad (véase la Tabla 6.8). Curiosamente, los índices de validez de la HAD en este análisis factorial fueron peores que los del BDI-FS, ya que tan sólo 11 de sus 14 ítems mostraron saturaciones iguales o mayores de 0,40 en sus respectivos factores (depresión o ansiedad) y saturaciones menores de 0,25 en el factor no correspondiente (ansiedad o depresión), mientras que en el caso del BDI-FS todos sus ítems se comportaron de la forma en que se esperaba, es decir, saturaciones significativas ($\geq 0,40$) en el primer factor de depresión y saturaciones despreciables ($< 0,25$) en el segundo factor de ansiedad (véase la Tabla 6.8).

6.3.7. Validez de criterio

Para determinar qué pacientes de la presente muestra padecían o no un trastorno depresivo mayor se tomó como criterio de referencia el diagnóstico

realizado por la investigadora (psicóloga clínica con experiencia) mediante la entrevista estructurada SCID-I. A partir de este criterio, se calcularon varios índices para valorar la capacidad o exactitud diagnóstica del BDI-FS para detectar pacientes con o sin trastorno depresivo mayor. Estos índices se calcularon a través de un análisis de la curva ROC o curva de rendimiento predictivo. Esta curva se construye mediante un gráfico de dos ejes donde, para cada puntuación del instrumento, se expresa su sensibilidad en el eje de la ordenada y el complementario de la especificidad en el eje de la abscisa. De esta forma se puede encontrar el punto de corte del BDI-FS que optimice tanto la sensibilidad como la especificidad del diagnóstico de trastorno depresivo mayor. Los puntos resultantes unidos por una línea dibujan una curva, y el área por debajo de dicha curva refleja la valoración global del rendimiento diagnóstico de la prueba o test en cuestión. Una curva ROC que coincida con la diagonal de los ejes de coordenadas representa la exactitud diagnóstica que se obtendría al azar (p. ej., lanzando una moneda al aire para diagnosticar un trastorno depresivo mayor en un paciente) e implica un área bajo la curva de 0,50. Por tanto, un área superior a 0,50 indica una exactitud diagnóstica por encima del azar y un área igual a 1 indica una exactitud perfecta, de manera que Swets (1988, p. 1292) considera que valores del área bajo la curva “entre 0,50 y 0,70 aproximadamente representan una exactitud bastante baja—la proporción de verdaderos positivos no es mucho mayor que la proporción de falsos positivos en cualquier punto a lo largo de la curva. Valores del [área bajo la curva] entre aproximadamente 0,70 y 0,90 representan exactitudes que son útiles para algunos propósitos, y valores más altos representan una exactitud bastante alta”. Los criterios de Swets (1988) para interpretar los valores del

área bajo la curva han sido adoptados por otros investigadores. Por ejemplo, tanto Fischer, Bachmann y Jaeschke (2003) como Streiner y Cairney (2007) consideran que un test con un área bajo la curva entre 0,50 y 0,70 tiene exactitud baja, entre 0,70 y 0,90 exactitud moderada, y más de 0,90 exactitud alta.

En la Figura 6.3 se muestra la curva ROC correspondiente a las puntuaciones del BDI-FS en la muestra de pacientes de AP de este estudio. El área bajo dicha curva ROC fue 0,97, con un intervalo de confianza al 95% de 0,92-0,99 y un error típico de 0,012, por lo que, según los valores convencionales anteriormente descritos, puede considerarse que la exactitud diagnóstica del BDI-FS para detectar personas con trastorno depresivo mayor ente los pacientes de AP fue alta.

Figura 6.4. Curva ROC o de rendimiento predictivo de las puntuaciones del BDI-FS en la muestra de pacientes de atención primaria del Estudio 1º

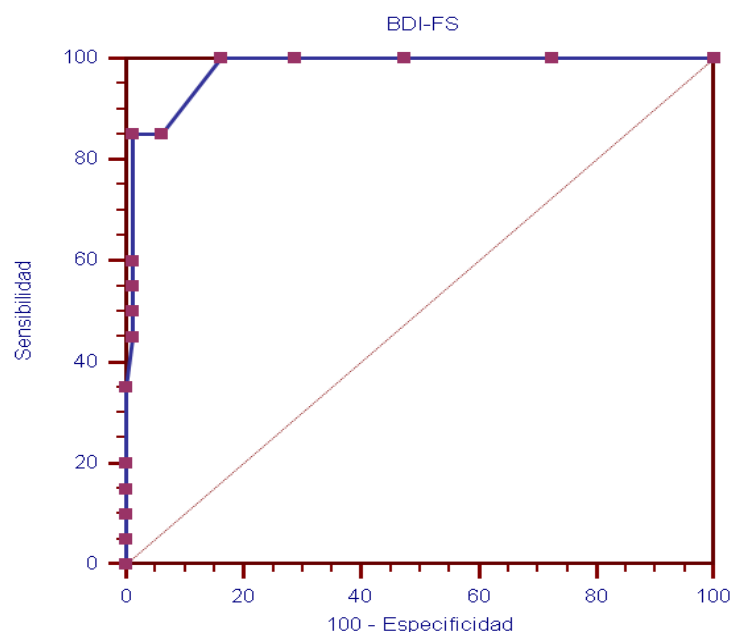


Tabla 6.9. Índices de rendimiento diagnóstico para los puntos de corte del BDI-FS que obtuvieron los mejores índices en la muestra de pacientes de atención primaria del Estudio 1º

Punto de corte*	Sensibilidad	Especificidad	Suponiendo que la prevalencia de la depresión = 20%				Suponiendo que la prevalencia de la depresión = 13,2%				Eficiencia	Índice de Youden	Kappa
			Valor predictivo positivo		Valor predictivo negativo		Valor predictivo positivo		Valor predictivo negativo				
			Valor	predictivo	Valor	predictivo	Valor	predictivo	Valor	predictivo			
			positivo	negativo	positivo	negativo	positivo	negativo	positivo	negativo			
3	100	71,25	46,5	100	34,6	100	0,77	0,71	0,50				
4	100	83,75	60,6	100	48,3	100	0,87	0,84	0,67				
5	85	93,75	77,3	96,2	67,4	97,6	0,92	0,79	0,76				
6	85	98,75	94,4	96,3	91,2	97,7	0,96	0,84	0,87				
7	60	98,75	92,3	90,8	88,0	94,2	0,91	0,59	0,68				

Nota. *Igual o mayor que esa puntuación.

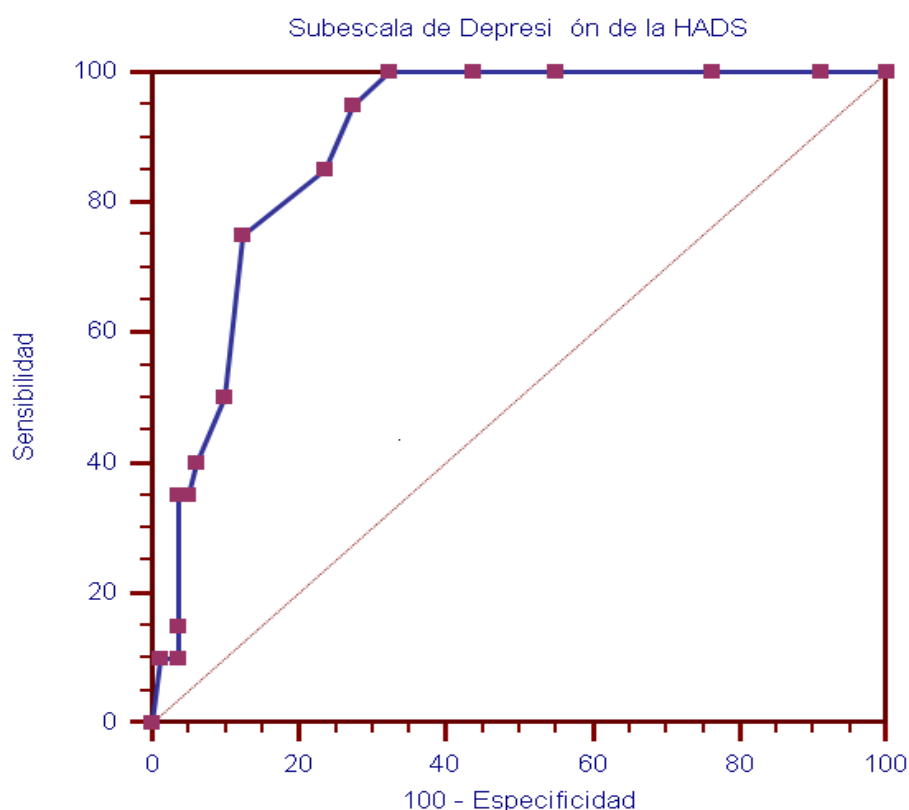
En la Tabla 6.9 se recogen para distintos puntos de corte del BDI-FS sus índices de sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo, eficiencia diagnóstica, índice de Youden y concordancia (kappa) con el diagnóstico de la SCID-I.

En estudios previos con el BDI-FS se ha encontrado que una puntuación igual o mayor de 4 (punto de corte = 4) obtiene los mejores resultados en los índices diagnósticos anteriormente mencionados y ha sido el punto de corte que más se ha utilizado (Beck et al., 1997, 2000; Steer, Cavalieri, Leonard y Beck, 1999). En el manual de la versión original del BDI-FS, se recomienda el uso de puntuaciones iguales o mayores al rango de 3 a 5 para obtener la máxima eficiencia clínica en detectar el trastorno depresivo mayor en poblaciones médicas en las cuales la prevalencia subyacente del trastorno depresivo mayor varíe de 5 a 40% (Beck et al., 2000, p. 24). Como puede verse en la Tabla 6.9, en el presente estudio el punto de corte de 4 obtuvo buenos índices de rendimiento diagnóstico (p. ej., un 100% de sensibilidad y un 84% de especificidad), al igual que las puntuaciones entre 3 y 5, pero ninguno de ellos fue el que mejores índices obtuvo. El punto de corte de 6 (una puntuación igual o mayor que 6) obtuvo los mejores resultados de rendimiento diagnóstico, con un 85% de sensibilidad y un 99% de especificidad. Así, cuando se comparan los índices más globales de rendimiento diagnóstico como, por ejemplo, el índice de eficiencia o el de Youden, se puede observar en la Tabla 6.9 que el índice de eficiencia fue mayor con el punto de corte de 6 que con el de 4 (0,96 frente a 0,87), aunque ambos puntos de corte presentaron el mismo índice de Youden (0,84). Es más, el punto de corte de 6 presentaba un mayor grado de acuerdo con el diagnóstico proporcionado por la SCID-I que el punto de corte

de 4 (0,87 frente a 0,67) y, así mismo, considerando conjuntamente los valores predictivos positivos y negativos, éstos eran superiores para el punto de corte de 6 en comparación a los del punto de corte de 4. Suponiendo una prevalencia del trastorno depresivo mayor en AP del 20% (la observada en el presente estudio), se obtuvieron para el punto de corte de 6 un valor predictivo positivo del 94,4% y un valor predictivo negativo de 96,3%, mientras que suponiendo una prevalencia del trastorno depresivo mayor en AP del 13,2% (la estimada en el Capítulo 1 a partir de los resultados de los estudios previos españoles), se obtuvieron un valor predictivo positivo del 91,2% y un valor predictivo negativo del 97,7%. En resumen, para el punto de corte de 6, todos los valores predictivos, positivos o negativos, superaron el valor de 90%, mientras que, tal y como se muestra en la Tabla 6.9, los valores predictivos positivos para el punto de corte de 4 fueron todos ellos menores de 70%, aunque sus valores predictivos negativos sí fueron mayores de 90%.

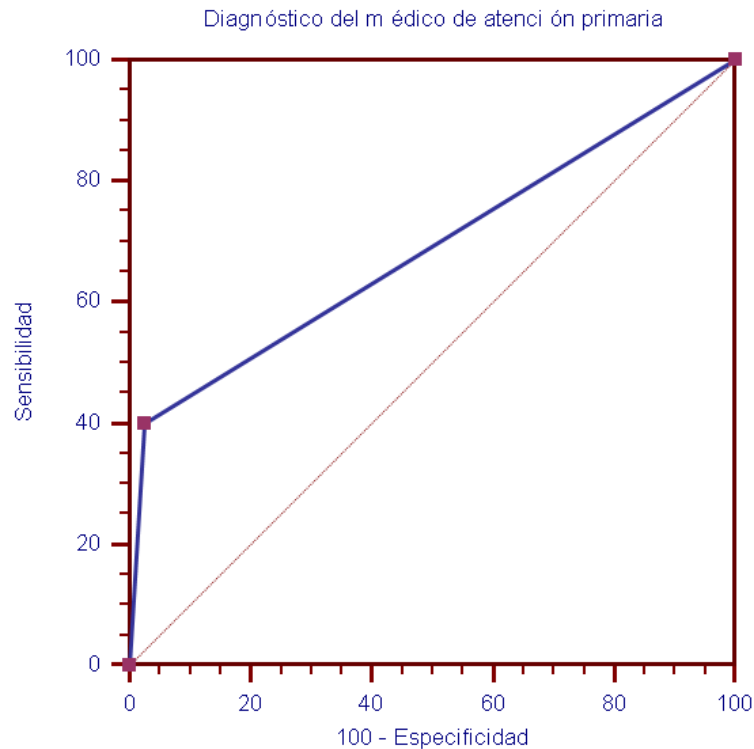
En cuanto al rendimiento diagnóstico de la subescala de depresión de la HAD, en la Figura 6.4 se presenta la curva ROC correspondiente a sus puntuaciones en la muestra del presente estudio. Para esta subescala, el área bajo la curva arrojó un valor de 0,89, con un intervalo de confianza al 95% de (0,814-0,945) y un error típico de 0,031. El punto de corte de la subescala de depresión de la HAD que mejores índices diagnósticos obtuvo fue una puntuación igual o mayor de 6 (punto de corte = 6), con una sensibilidad igual a 95%, una especificidad igual a 72,5%, un *kappa* de concordancia igual a 0,48, un índice de eficiencia diagnóstica igual a 0,77 y un índice de Youden igual a 0,67.

Figura 6.5. Curva ROC o de rendimiento predictivo de las puntuaciones de la subescala de depresión de la HAD en la muestra de pacientes de atención primaria del Estudio 1º



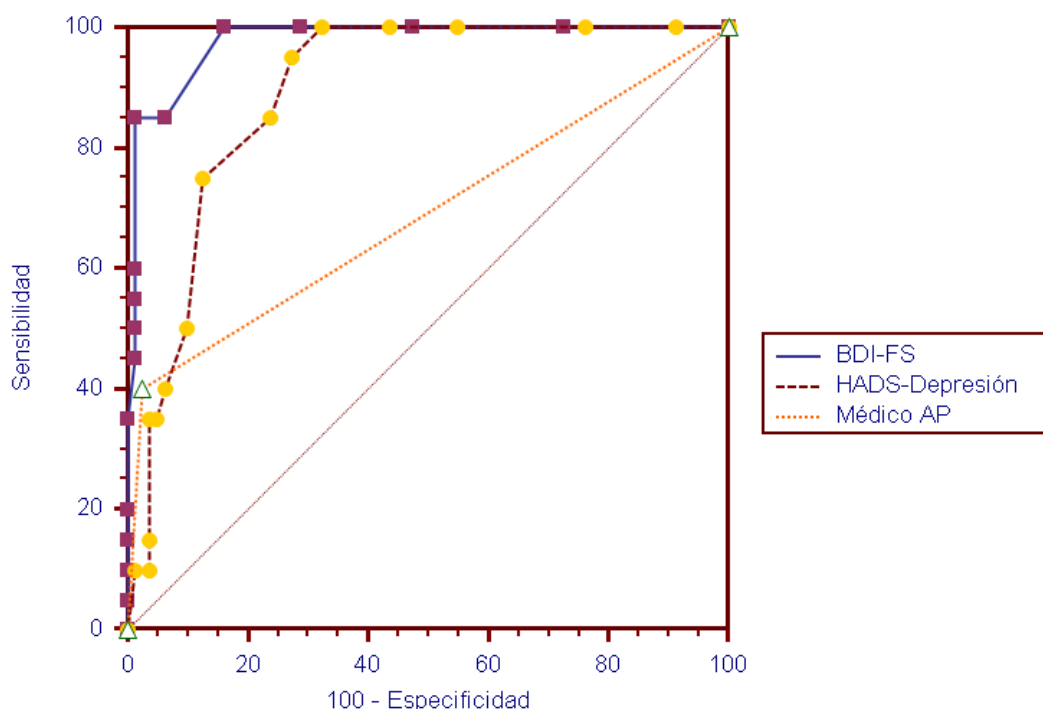
En la Figura 6.5 se muestra la curva ROC del juicio clínico sobre la existencia o no de un trastorno depresivo realizado por los médicos de AP que participaron en esta investigación frente al criterio del diagnóstico de trastorno depresivo mayor realizado por un especialista a partir de una entrevista diagnóstica estructurada. El área bajo la curva ROC para el juicio clínico del MAP fue igual a 0,68, con un intervalo de confianza al 95% de 0,81-0,94 y un error típico de 0,056. Los otros índices de rendimiento diagnóstico del juicio clínico del MAP arrojaron los siguientes valores: una sensibilidad del 40%, una especificidad del 97,5%, un valor predictivo positivo del 80%, un valor predictivo negativo del 86,6%, un índice de eficiencia diagnóstica igual a 0,86, un índice de Youden igual a 0,38, y un índice *kappa* de concordancia igual a 0.46.

Figura 6.6. Curva ROC o de rendimiento predictivo del diagnóstico del médico de atención primaria en la muestra de pacientes de atención primaria del Estudio 1º



Al comparar las curvas ROC del BDI-FS, de la subescala de depresión de la HAD y del juicio clínico del médico de AP, se pudo apreciar visualmente el mejor rendimiento diagnóstico del BDI-FS frente a los otros procedimientos, ya que la curva de este último instrumento era la que más cerca estaba del vértice superior izquierda del gráfico y que representa los valores más altos de sensibilidad y de especificidad (véase la Figura 6.6).

Figura 6.7. Comparación de las curvas ROC del BDI-FS, de la subescala de depresión de la HAD y del diagnóstico del médico de atención primaria



La impresión visual que se desprende de observar la Figura 6.6 quedó confirmada estadísticamente al comparar las áreas bajo las curvas ROC correspondientes. La prueba estadística realizada según el método de DeLong et al. (1988) para examinar la significación estadística de la diferencia entre las áreas bajo la curva ROC del BDI-FS y el juicio médico (0,97 frente a 0,68) indicó que el BDI-FS detectaba significativamente mejor a los pacientes con trastornos depresivos que el criterio médico ($Z = 5,03$, $p < 0,0001$). De igual manera, el área bajo la curva ROC del BDI-FS fue significativamente mayor que el área bajo la curva ROC de la subescala de depresión de la HAD (0,97 frente a 0,89; $Z = 2,94$, $p < 0,003$), indicando que el BDI-FS discriminaba significativamente mejor a los pacientes de AP con trastornos depresivos que la

subescala de depresión de la HAD. A su vez, el área bajo la curva ROC de la subescala de depresión de la HAD fue significativamente mayor que el área bajo la curva ROC del juicio médico (0,89 frente a 0,68; $Z = 3,007$, $p < 0,002$), lo que implicaba que dicha subescala detectaba con mayor exactitud a los pacientes con trastornos depresivos que el médico de AP.

Finalmente, al comparar el resto de indicadores de rendimiento diagnósticos que se han calculado para el BDI-FS, la subescala de depresión de la HAD y el juicio del médico de AP (sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo, eficiencia, índice de Youden y *kappa*), se puede observar que, en todos y cada uno de esos indicadores, el mejor punto de corte del BDI-FS mostraba un rendimiento diagnóstico mejor que el que ofrecía el criterio médico o el mejor punto de corte de la subescala de depresión de la HAD (véase la Tabla 6.10).

Tabla 6.10. Índices de rendimiento diagnóstico para el juicio clínico del médico de atención primaria (AP) y para el mejor punto de corte del BDI-FS y de la subescala de depresión de la HAD (HAD-D)

Instrumento (punto de corte)	Sensibilidad	Especificidad	Suponiendo que la prevalencia de la depresión = 20%				Suponiendo que la prevalencia de la depresión = 13,2%				Eficiencia	Índice de Youden	Kappa
			Valor		Valor		Valor		Valor				
			predictivo		predictivo		predictivo		predictivo				
			positivo	negativo	positivo	negativo	positivo	negativo	positivo	negativo			
BDI-FS (6)	85	98,75	94,4	96,3	91,2	97,7	91,2	97,7	91,2	97,7	0,96	0,84	0,87
HAD-D (6)	95	72,5	46,3	98,3	34,4	99	34,4	99	34,4	99	0,77	0,67	0,48
Médico de AP	40	97,5	80	86,7	70,9	91,4	70,9	91,4	70,9	91,4	0,86	0,37	0,46

6.4. Discusión

A partir de los resultados encontrados en este primer estudio se puede concluir que el BDI-FS muestra unas buenas propiedades psicométricas como instrumento para la evaluación de la sintomatología depresiva en pacientes españoles de AP y para la detección de trastornos depresivos en dichos pacientes.

6.4.1. Validez factorial del BDI-FS

Coincidiendo con la hipótesis inicial de este estudio y con la estructura factorial del BDI-FS original (Beck et al. 2000), la versión española del BDI-FS mostró en una muestra de pacientes españoles de AP una solución factorial claramente unifactorial que reflejaba una dimensión cognitivo-afectiva de la depresión, es decir, sintomatología depresiva cognitivo-afectiva, lo cual era previsible teniendo en cuenta que el BDI-FS procede de los ítems del BDI-II que *a priori* evalúan síntomas cognitivos y afectivos y que, consistentemente, forman una dimensión cognitivo-afectiva en los propios análisis factoriales del BDI-II (Sanz, Navarro y Vázquez, 2003; Sanz, Perdigón y Vázquez, 2003; Sanz et al., 2005). De esta manera, el BDI-FS cubriría los criterios más puramente psicológicos de la depresión que, a su vez, pueden pasar más fácilmente desapercibidos para el médico de AP en detrimento de los somáticos.

6.4.2. Sensibilidad del BDI-FS

De igual manera a lo encontrado en estudios previos (p. ej., Beck et al., 2000), las puntuaciones del BDI-FS en la presente muestra de pacientes de AP no se distribuyeron de forma normal. Aunque esto podría indicar un problema del BDI-FS para evaluar diferencias individuales en sintomatología depresiva, dicho resultado es coherente con la naturaleza del BDI-FS y es una característica que cabría esperar del instrumento cuando se utiliza en muestras de pacientes de AP. El BDI-FS se desarrolló a partir del BDI-II y éste es un test diseñado principalmente como “un instrumento de evaluación de la gravedad de la depresión en pacientes adultos y adolescentes con diagnóstico psiquiátrico” (Beck et al., 1996, p. 6), por lo que sus ítems presentan alternativas de respuesta que tratan de apresar toda la amplitud de la gravedad sintomatológica depresiva que pueden presentar los pacientes con trastornos mentales que, lógicamente, es mucho mayor que la que pueden presentar, por ejemplo, los pacientes de AP, la mayoría de los cuales tienen enfermedades médicas, pero no trastornos mentales. En consecuencia, las puntuaciones del BDI-II se suelen distribuir de forma normal cuando el instrumento se aplica en muestras de pacientes con trastornos psicopatológicos, pero no en otros tipos de muestras en los que cabría esperar una menor frecuencia y gravedad de la sintomatología depresiva (Beck et al., 1996). Por ejemplo, la adaptación española del BDI-II presenta una distribución normal de sus puntuaciones en una muestra de pacientes ambulatorios con trastornos mentales (Sanz et al., 2005), mientras que presenta distribuciones parecidas a la encontrada con el BDI-FS en este estudio con pacientes de AP, es decir, distribuciones leptocúrticas desplazadas a la izquierda, en muestras de estudiantes

universitarios (Sanz, Navarro y Vázquez, 2003) y de adultos de la población general (Sanz, Perdigón y Vázquez, 2003). Algo parecido ocurre con las versiones previas del BDI, es decir, la versión original (BDI) y su versión revisada (BDI-IA) (Kendall, Hollon, Beck, Hammen e Ingram, 1987). Así, la adaptación española del BDI-IA también presenta una distribución normal de sus puntuaciones en muestras de pacientes ambulatorios con trastornos mentales (Vázquez y Sanz, 1999), mientras que presenta distribuciones leptocúrticas desplazadas a la izquierda en muestras de estudiantes universitarios (Sanz y Vázquez, 1998) y de adultos de la población general (Vázquez y Sanz, 1997).

En conclusión, el hecho de que las puntuaciones del BDI-FS en la muestra de pacientes de AP del presente estudio no se hayan distribuido de forma normal era esperable dadas las características del instrumento y de la población, y no representa un problema de sensibilidad del BDI-FS para detectar diferencias individuales en sintomatología depresiva entre los pacientes de AP ya que, por un lado, en la submuestra de pacientes con trastorno depresivo mayor sí que se encontró una distribución normal de las puntuaciones del BDI-FS y, por otro lado, los estadísticos de tendencia central (media) y dispersión (desviación típica) encontrados tanto en la muestra total de pacientes de AP como en la submuestra de pacientes con trastorno depresivo mayor se asemejan a los encontrados en muestras de pacientes de AP de otros países.

Así, en la muestra total de pacientes de AP del presente estudio se obtuvo en el BDI-FS una puntuación media igual a 2,85 y una desviación típica de 3,47, valores similares a los obtenidos por Steer y Thomas (1999) en una

muestra estadounidense de pacientes ambulatorios de medicina interna (media = 2,18 y $DT = 2,96$; véase también Beck et al., 2000) y a los obtenidos por Beck et al. (2000) en una muestra estadounidense de pacientes de AP (media = 3,23 y $DT = 3,57$). Si bien en otros estudios se han encontrado medias mayores como, por ejemplo, en el estudio de Beck et al. (1997) con 50 pacientes hospitalizados (media = 5,76 y $DT = 4,46$; véase también Beck et al., 2000) o en el estudio de Poole et al. (2009) con 1227 pacientes con dolor crónico (media = 7,11 y $DT = 4,30$), cabe reseñar la mayor prevalencia y gravedad de los cuadros depresivos en los pacientes que o bien sufren enfermedades tan graves que requieren hospitalización o bien sufren enfermedades que son crónicas, en comparación a la mayor diversidad, heterogeneidad y levedad de las enfermedades y síntomas depresivos que presentan los pacientes procedentes de consultas ambulatorias de AP. Por otro lado, en la submuestra de pacientes de AP con trastorno depresivo mayor del Estudio 1º se obtuvo una puntuación media en el BDI-FS igual a 9 ($DT = 4,2$), ligeramente superior a la puntuación media de los pacientes con trastorno depresivo mayor del estudio de Beck et al. (2000) con pacientes estadounidense de AP (media = 7,6 y $DT = 3,8$) y del estudio de Beck et al. (1997; véase también Beck et al., 2000) con pacientes ambulatorios estadounidenses de medicina interna (media = 6,5 y $DT = 2,5$).

6.4.3. Fiabilidad del BDI-FS

En relación con la fiabilidad del BDI-FS, entendida ésta como consistencia interna, los resultados del presente estudio tanto en términos del coeficiente *alfa* de Cronbach (igual a 0,85) como en términos de la correlación

media entre ítems (igual a 0,45) sugieren que la versión española del BDI-FS muestra buenos índices de fiabilidad en pacientes de AP de acuerdo a los estándares actuales (Bollen y Lennox, 1991; Briggs y Cheek, 1986; Prieto y Muñiz, 2000).

Esos índices, además, son acordes con los encontrados en un buen número de estudios con muy diversas muestras de pacientes médicos de diferentes países. Así, Steer et al. (1999) encontraron un *alfa* de 0,85 en una muestra de pacientes médicos adultos ambulatorios; Winter et al. (1999) encontraron un *alfa* de 0,88 con pacientes ambulatorios adolescentes con patología médica; Beck et al. (1999) encontraron un *alfa* de 0,86 en una muestra de pacientes médicos adultos hospitalizados; Kefetz et al. (2004) encontraron un *alfa* de 0,84 en pacientes con VIH; Scheinthal et al. (2001) encontraron un *alfa* de 0,83 en pacientes geriátricos; Golden et al. (2007) obtuvieron un *alfa* de 0,85 en pacientes con hepatitis C, y, en España, Orive et al. (2010) hallaron un *alfa* de 0,86 en pacientes con patologías médicas diversas. El coeficiente de fiabilidad más bajo encontrado en la literatura científica es el informado por Lam, Contreras, Mori, Gil et al. (2009) en población gestante peruana (*alfa* = 0,68), coeficiente que se aleja bastante del rango de 0,83-0,88 que informan los otros estudios anteriormente mencionados o el presente estudio. El motivo de ese bajo índice de fiabilidad de consistencia interna de Lam, Contreras, Mori, Gil et al. (2009) no está claro; por un lado, podría ser que el BDI-FS no fuera apropiado para pacientes gestantes, pero, por otro lado, cabe la posibilidad de que ese bajo índice fuera debido a alguna deficiencia metodológica del estudio en cuestión. En este sentido, por ejemplo, no queda claro en el estudio de Lam, Contreras, Mori, Gil et al. (2009) qué

versión española se utilizó o, si fue una versión española creada *ad hoc* para el estudio, cómo se realizó la traducción del BDI-FS al español y cómo se comprobó su legibilidad en la muestra de pacientes gestantes. Es más, parece que en dicho estudio realmente se aplicó el BDI-II y se analizaron los ítems correspondientes al BDI-FS (Lam, Contreras, Mori, Cuesta et al., 2009), de manera que las propiedades psicométricas del BDI-FS podrían haber sido distintas de haberse aplicado sus ítems en solitario, no como parte del BDI-II. Además, Lam, Contreras, Mori, Gil et al. (2009) informaron de una solución bifactorial del BDI-FS, solución muy poco habitual en la literatura científica sobre el BDI-FS, pero que podía responder también a deficiencias en la traducción o a deficiencias en la metodología del análisis factorial (p. ej., no se informa de los criterios utilizados para retener dos factores; se usó una rotación ortogonal *varimax* para la rotación de factores cuando lo apropiado hubiera sido una rotación oblicua dado que se presupone que los síntomas depresivos están relacionados).

Por otro lado, en el presente estudio todos los ítems del BDI-FS mostraron índices de consistencia interna adecuados (correlaciones corregidas ítem-total por encima de 0,30), destacando, en este sentido, los ítems de tristeza y de pérdida de placer, que mostraron las mayores correlaciones corregidas ítem-total ($r = 0,66$) y que, precisamente, son los síntomas cardinales en el diagnóstico de los trastornos depresivos según el DSM-IV (APA, 1994).

En relación con las correlaciones corregidas ítem-total, Golden et al. (2007) han abogado por la eliminación del ítem de Suicidio del BDI-FS ya que, en su estudio, la correlación corregida ítem-total de dicho ítem fue menor de

0,40 y su eliminación no alteró la fiabilidad del BDI-FS, y, en consecuencia, han construido con el resto de los ítems el denominado BDI-FS6 (en el Capítulo 4 se pueden consultar las propiedades psicométricas del BDI-FS6).

En una línea parecida, Wilhelm, Kotze, Waterhouse, Hadzi - Pavlovic y Parker (2004) consideran que el ítem de suicidio es una pregunta demasiado “directa o fuerte” para personas que están enfermas o en un proceso de afrontamiento de la muerte, por lo que consideran que su evaluación se debe realizar en una entrevista clínica más exhaustiva.

En el presente estudio se encontró una correlación corregida ítem-total para el ítem de Suicidio igual a 0,58, suficientemente grande como para no justificar su eliminación. De forma coincidente, Poole et al. (2009) encontraron una correlación corregida ítem-total grande para el ítem de Suicidio ($r = 0,55$) en una muestra de 1227 personas con dolor crónico, y concluyen que la sintomatología suicida es de suficiente importancia para ser evaluada desde el principio tanto a nivel de cribado como en una entrevista clínica posterior. Además, lejos de incrementar el riesgo de suicidio, el poder reconocer y expresar los deseos de muerte facilita que no se llegue a consumir el suicidio (Murphy, 1986; Doyle, 1990; Bobes, Saíz, García-Portilla, Bascarán y Bousoño, 2004). En la misma dirección, Joiner et al. (2005), en su estudio sobre la evaluación de la depresión en adultos basada en pruebas, consideran que no sólo hay que evaluar con cuestionarios la ideación y los deseos de muerte, sino también la preparación y los planes de suicidio.

En conclusión, los resultados del presente estudio no parecen aconsejar la eliminación del ítem de Suicidio del BDI-FS ya que no sólo obtuvo un índice

de consistencia interna adecuado, sino que en la solución unifactorial del BDI-FS obtuvo una saturación factorial de 0,63, refrendando así su validez y su necesaria inclusión en el instrumento.

6.4.4. Validez convergente del BDI-FS

Los resultados del presente estudio demuestran que en pacientes de AP el BDI-FS correlaciona de manera importante ($> 0,70$) con otras medidas de depresión, tanto heteroevaluadas (la medida de episodio depresivo mayor obtenida a partir de la SCID-I) como autoevaluadas (la subescala de depresión de la HAD), lo que permite concluir que la versión española del BDI-FS tiene unos índices de validez convergente excelentes de acuerdo a los estándares actuales ($r \geq 0,60$; Prieto y Muñiz, 2000) y, por tanto, que es un instrumento que mide fundamentalmente depresión, al menos, en la misma medida o mejor que otros instrumentos de cribado habitualmente utilizados en la práctica clínica con pacientes de AP.

Además, los índices de validez convergente encontrados son coherentes con los encontrados en la literatura científica con muestras diversas de pacientes médicos. Por ejemplo, Beck, Steer, Ball, Ciervo y Kabat (1997) hallaron una correlación entre el BDI-FS y el diagnóstico de trastorno del estado de ánimo del PRIME-MD (*Primary Care Evaluation of Mental Disorders*) igual a 0,69 en pacientes de AP; Beck, Guth, Steer y Ball (1997) encontraron una correlación entre el BDI-FS y la subescala de depresión de la HAD igual a 0,61 en una muestra de pacientes médicos hospitalizados; Scheinthal et al. (2001) encontraron una correlación del BD-FS con el GDS (*Geriatric*

Depression Scale) igual a 0,81 en población geriátrica, y, en España, Orive et al. (2010) hallaron en una muestra de pacientes con diversas patologías médicas una correlación del BDI-FS con el DMI-18 (*Depression in the Medically Ill-18*) igual a 0,88 y con el DMI-10 (*Depression in the Medically Ill-10*) igual a 0,87.

6.4.5. Validez discriminante del BDI-FS

En el presente estudio, la versión española del BDI-FS mostró una correlación positiva, significativa y grande ($> 0,50$) con la subescala de ansiedad del HAD, con un valor parecido al encontrado en estudios similares (0,65 en este estudio frente a 0,62 en el estudio de Beck et al., 1999).

Esta correlación podría, en principio, sugerir falta de validez discriminante por parte del BDI-FS cuando se utiliza en muestras de pacientes de AP. No obstante, dado el solapamiento de las definiciones de ansiedad y depresión (Sanz y Navarro, 2003), cierta correlación entre sus medidas es esperable y, de hecho, se encuentra habitualmente tanto en muestras de pacientes psicopatológicos como en muestras de la población general (Sanz y Navarro, 2003; Magán et al., 2008; Sanz et al., 2011). Por tanto, la cuestión debería plantearse en términos de si el BDI-FS, a pesar de su alta correlación con medidas de ansiedad como la subescala de ansiedad de la HAD, (a) correlaciona más con medidas de depresión que con medidas de ansiedad, y (b) permite medir una sintomatología afectiva diferente de la ansiedad.

Respecto al primer punto, los resultados de este estudio indican que las correlaciones que presentó el BDI-FS con la medida de episodio depresivo

mayor de la SCID-I y con la subescala de depresión de la HAD fueron mayores que la correlación que presentó el BDI-FS con la subescala de ansiedad de la HAD, y en el caso de la correlación del BDI-FS con la subescala de depresión de la HAD, tal diferencia fue estadísticamente significativa.

En relación con el segundo punto anteriormente mencionado, los resultados del análisis factorial conjunto de los ítems del BDI-FS y de la HAD revelan que los ítems del BDI-FS y de la subescala de ansiedad de la HAD son distinguibles entre sí y forman dos factores distintos, uno de depresión y otro de ansiedad, respectivamente.

Por lo tanto, estos últimos resultados sugieren que la correlación entre el BDI-FS y la subescala de ansiedad de la HAD en pacientes de AP puede deberse más a la relación entre los constructos de ansiedad y depresión, que a un problema de falta de validez discriminante por parte del BDI-FS.

Por otro lado, en este estudio se han obtenido datos sobre la validez discriminante de la subescala de depresión de la HAD que son coincidentes con los problemas que dicha escala ha mostrado al respecto en la literatura científica (p. ej., Herrero et al., 2003). En concreto, en la muestra de pacientes de AP de la presente investigación, la correlación de la subescala de depresión de la HAD con la subescala de ansiedad de la HAD fue de 0,66, pero, al contrario de lo que ocurría con el BDI-FS, dicha correlación no sólo no fue inferior a la correlación de la subescala de depresión de la HAD con la medida de episodio depresivo mayor de la SCID-I, sino que fue más alta (0,66 frente a 0,55).

De forma consistente, los índices de validez de la subescala de depresión de la HAD en el análisis factorial conjunto de los ítems del BDI-FS y de la HAD fueron peores que los del BDI-FS e incidían en los problemas de validez discriminante mencionados anteriormente, ya que uno de sus 7 ítems (el ítem 14: Capacidad para disfrutar con un buen libro, programa de radio o televisión) no alcanzó a saturar más de 0,25 en el factor de depresión y, en cambio, saturaba más de 0,25 en el factor de ansiedad y, así mismo, otro de sus ítems (el ítem 6: Alegría), aunque mostraba una saturación igual o mayor de 0,40 en el factor de depresión, también saturaba más de 0,25 en el factor de ansiedad.

Estos resultados son coherentes con los resultados de estudios previos como, por ejemplo, el de Herrero et al. (2003), quienes encontraron que la subescala de ansiedad de la HAD correlacionaba 0,71 y 0,68 con las escalas de ansiedad estado y rasgo, respectivamente, del STAI (Spielberger et al., 1988), pero que la subescala de depresión de la HAD mostraba correlaciones ligeramente superiores con esas dos escalas de ansiedad del STAI (0,74 en ambos casos). Es más, estos problemas de validez discriminante frente a la ansiedad de la subescala de depresión de la HAD quizás habría que ponerlos en el contexto de algunos resultados curiosos sobre su rendimiento como instrumento de cribado de los trastornos depresivos como, por ejemplo, los encontrados por Sultan et al. (2010) que indicaban que la puntuación total de la HAD, es decir, la puntuación resultante de sumar los ítems de sus dos subescalas (depresión y ansiedad), mostraba mejores índices de rendimiento diagnóstico para detectar trastornos depresivos en pacientes con diabetes que la propia puntuación de la subescala de depresión de la HAD.

6.4.6. Validez de criterio del BDI-FS: rendimiento diagnóstico

Una de las propiedades psicométricas más importantes para avalar la utilización de la versión española BDI-FS como instrumento de cribado o despistaje en la detección de trastornos depresivos en pacientes de AP es su validez de criterio en términos de su rendimiento diagnóstico respecto a un criterio de referencia sobre la presencia de un trastorno depresivo mayor. En este estudio, el criterio elegido es el que actualmente se considera el “patrón oro” en el campo del psicodiagnóstico: el diagnóstico realizado por un clínico con experiencia utilizando una entrevista diagnóstica estructurada. En concreto, en la presente investigación, el diagnóstico realizado por una psicóloga clínica con experiencia utilizando el módulo de trastornos del estado de ánimo de la SCID-I.

Con ese criterio de referencia, los resultados obtenidos en relación con el rendimiento diagnóstico de la versión española del BDI-FS son muy buenos. Por ejemplo, para el BDI-FS se obtuvo un área bajo la curva ROC igual a 0,97. Este estadístico es un índice global de rendimiento predictivo o diagnóstico y, según los estándares actuales (Fischer et al., 2003; Streiner y Cairney, 2007; Swets, 1988), se puede considerar que la exactitud diagnóstica del BDI-FS para detectar el trastorno depresivo mayor es alta ($> 0,90$).

El valor de ese índice, además, es consistente con los encontrados en estudios previos realizados con el BDI-FS en muestras de pacientes médicos. Por ejemplo, Winter et al. (1999) informaron de un área bajo la curva ROC de 0,98, Beck et al. (1999) de un área de 0,92, Steer et al. (1999) de un área de

0,99, y Poole et al. (2009) de un área de 0,94. Fuera del rango de exactitud alta, pero cerca de los valores superiores del rango de exactitud moderada, se encuentran los valores del área bajo la curva ROC informados por Golden et al. (2007) y, en España, por Orive et al. (2010), que en ambos estudios fueron de 0,85.

En definitiva, los resultados de este estudio unido a los resultados de esos estudios previos, permiten concluir que el BDI-FS tiene un rendimiento diagnóstico bueno o alto para detectar pacientes médicos con trastorno depresivo mayor.

Es más, los resultados del presente estudio también sugieren que el rendimiento diagnóstico de la versión española del BDI-FS en pacientes de AP es significativamente mayor que el de la subescala de depresión de la HAD o que el del juicio clínico del médico de AP.

Efectivamente, aunque el área bajo la curva ROC para la subescala de depresión de la HAD fue en este estudio (0,89) mayor que las encontradas para dicho instrumento por Beck y Steer (1999), Golden et al. (2007) y Orive et al. (2010) (0,89 frente a 0,74, 0,78 y 0,84, respectivamente), dicho valor fue significativamente menor que el del BDI-FS.

Respecto al juicio clínico del médico de AP, los resultados de este estudio también indicaron que su área bajo la curva ROC fue significativamente menor que el del BDI-FS, por lo que se podría concluir que la utilización del BDI-FS por parte de los médicos españoles de AP probablemente supondría un aumento importante en la precisión con que podrían detectar y diagnosticar un trastorno depresivo entre sus pacientes.

En cuanto a la cuestión de cuál es el mejor punto de corte del BDI-FS para detectar pacientes de AP que sufren trastorno depresivo mayor, Beck et al. (2000, p. 24) propusieron que las puntuaciones en el rango de 3-5 ofrecían los mejores índices de rendimiento diagnóstico en poblaciones médicas en las cuales la prevalencia del trastorno depresivo mayor estuviera entre 5 y 40% y, de hecho, la puntuación de 4 había sido una de las que mejores índices de sensibilidad y especificidad había obtenidos en estudios previos (Beck et al., 1997, 2000; Steer, Cavalieri, Leonard y Beck, 1999). En el presente estudio, sin embargo, con una prevalencia del trastorno depresivo mayor del 20%, el punto de corte de 6, es decir, una puntuación igual o mayor que 6 en el BDI-FS, fue el valor que obtuvo los mejores índices diagnósticos, con un 85% de sensibilidad y un 99% de especificidad, con valores predictivos positivos y negativos por encima de 90%, y con un índice kappa de concordancia con el diagnóstico de referencia de 0,87. Todos estos índices no sólo fueron buenos o excelentes de acuerdo a los estándares actuales (Cicchetti, 1994), sino que, en general, fueron superiores a los índices de los demás puntos de corte del BDI-FS, a los índices del mejor punto de corte de la subescala de depresión de la HAD y a los índices del juicio diagnóstico médico.

Aunque con índices diagnósticos peores a los del presente estudio, los resultados del estudio de Golden et al. (2007) también indicaban que el punto de corte de 6 era mejor que el de 4. En concreto, Golden et al. (2007) encontraron en una muestra de pacientes con diabetes que, frente al punto de corte de 4, el punto de corte de 6 mejoraba de manera estadísticamente significativa la predicción de un trastorno depresivo mayor ($p = 0,021$), presentando una sensibilidad del 84%, una especificidad del 76%, un valor

predictivo positivo de 58%, un valor predictivo negativo de 91%, y un kappa de 0,53 de concordancia con el diagnóstico de la SCID-I.

Por último, merece la pena destacar que el punto de corte de 6 del BDI-FS obtuvo en este estudio un valor predictivo positivo de 94,4%, suponiendo una prevalencia del trastorno depresivo mayor en AP del 20% (la observada en el presente estudio), mientras que obtuvo un valor predictivo positivo de 91,2%, suponiendo una prevalencia del trastorno depresivo mayor en AP del 13,2% (la estimada en el Capítulo 1 a partir de los resultados de los estudios españoles previos). Este índice es uno de los más útiles en la práctica clínica ya que permite responder a una de las preguntas más frecuentes en la misma: ¿cuál es la probabilidad de que este paciente tenga un trastorno depresivo mayor ahora que ya sé que tiene una puntuación, por ejemplo, de 6 o más? Por tanto, según los resultados del presente estudio, se podría estimar que cuando un paciente español de AP obtiene una puntuación de 6 o más en el BDI-FS, la probabilidad de que sufra un trastorno depresivo mayor oscila entre el 91 y el 94%, probabilidad que sin duda supone una garantía para que el médico de AP pueda detectar a dichos pacientes y evaluar y tratar adecuadamente su depresión.

En resumen, los resultados de este estudio señalan que, como instrumento de cribado, el BDI-FS tiene buenos índices de validez de criterio y es una alternativa mejor, en términos de esos índices, que otros instrumentos de cribado de la depresión frecuentemente utilizados en España como, por ejemplo, la subescala de la depresión de la HAD, y mejor que el juicio diagnóstico médico sin la ayuda de ningún instrumento de evaluación o cribado de la depresión.

6.4.7. Limitaciones del Estudio 1º

En general, los resultados del presente estudio permiten concluir que la versión española del BDI-FS tiene buenas propiedades psicométricas en pacientes médicos de AP y que es un instrumento prometedor para la detección del trastorno depresivo mayor en dichos pacientes. Sin embargo, esta conclusión general como las conclusiones específicas sobre sus propiedades psicométricas a las que se ha llegado en los epígrafes anteriores deberían considerarse en el contexto de algunas de las limitaciones que tiene el presente estudio.

En primer lugar, la muestra de pacientes de AP evaluados fue una muestra incidental o de conveniencia procedente de un centro de atención primaria de una ciudad española y remitidos por un grupo de conveniencia de médicos, sin que se hubieran seguido en su selección criterios de muestreo aleatorio respecto a otros centros de atención primaria de esa ciudad, respecto a otras ciudades españolas o respecto a otros médicos de ese centro de atención primaria. Por todo ello, cabría cuestionarse la validez externa o representatividad de los resultados encontrados. En este sentido, por ejemplo, en la presente muestra de pacientes con AP se detectó a partir de la SCID-I un 20% de pacientes con trastorno depresivo mayor, porcentaje que si bien coincide con la prevalencia de ese trastorno en pacientes de AP que se ha estimado en algún estudio español previo (p. ej., Gabarrón et al., 2002), es ligeramente superior a la prevalencia estimada en otros y que oscila entre 9,6 y 16,8% (Aragonés et al., 2001, 2004, 2005; Caballero et al., 2008; Ferrer y Rodríguez, 1999; Serrano et al., 2010). No obstante, son reseñables las diferencias existentes entre esos estudios en cuanto a los instrumentos de

evaluación, criterios de referencia, tipo de muestreo, tamaño muestral, composición demográfica de la muestra, etc., de forma que tales diferencias podrían explicar las discrepancias entre los resultados de prevalencia de la depresión sin necesidad de suponer que la muestra de este estudio es especialmente poco representativa de la población española de pacientes de AP.

Futuras investigaciones sobre la versión española del BDI-FS deberían tratar de solventar esa limitación utilizando procedimientos de muestreo de pacientes de AP más adecuados. En cualquier caso, es obvio que una replicación de los resultados del presente estudio con una muestra distinta de pacientes de AP también podría ayudar a garantizar su validez externa y, precisamente, este será uno de los objetivos del Estudio 2º de esta tesis doctoral.

En segundo lugar, aunque el diagnóstico de referencia de los pacientes se fundamentó en el juicio de una psicóloga clínica con experiencia sobre la base de una entrevista diagnóstica estructurada, no fue posible examinar la fiabilidad de dichos diagnósticos. Futuras investigaciones sobre la versión española del BDI-FS, sobre todo sobre su rendimiento diagnóstico, deberían tratar también de solucionar esta limitación.

En suma, con las cautelas anteriormente citadas, se puede concluir que la versión española del BDI-FS parece tener unas propiedades psicométricas aceptables como instrumento de evaluación de la sintomatología depresiva en pacientes de AP y como instrumento de cribado o despistaje de trastornos depresivos en esos pacientes, propiedades que recomiendan su uso en este tipo de población y con ese propósito. Sin embargo, es obvio que la validación

de cualquier instrumento es un proceso sin fin que requiere el análisis del mayor número de propiedades psicométricas y en el mayor número de condiciones, análisis que, a la postre, sustente las muy diversas interpretaciones y usos que se pueden proponer de las puntuaciones del test. En este sentido, en el presente estudio no se ha abordado una propiedad psicométrica básica como es la fiabilidad de estabilidad temporal o fiabilidad test-retest, o la validez convergente-discriminante se ha centrado en únicamente dos instrumentos de medida de la depresión y la ansiedad, la HAD y la SCID-I, cuando existe una gran variedad de instrumentos ya validados que también miden dichos constructos y que presentan características y fortalezas distintas a las de aquéllos y que, por tanto, permitirían explorar la validez convergente-discriminante del BDI-FS con mayor profundidad. Por tanto, con el fin de seguir explorando las bondades del BDI-FS y poder aportar información complementaria y más exhaustiva que no ha sido considerada en este estudio, se decidió realizar el Estudio 2º que se describe en el siguiente capítulo.

Capítulo 7

Segundo Estudio: Ampliación y Replicación de las Propiedades Psicométricas del BDI-FS en Atención Primaria en España

7.1. Introducción

A raíz de los buenos resultados encontrados en el Estudio 1º de esta tesis doctoral, parece que hay indicios para sugerir que el BDI-FS puede ser también en España un buen instrumento en AP para evaluar sintomatología depresiva y para el cribado o despistaje de pacientes que sufren trastornos depresivos y, por tanto, parece interesante proseguir con la investigación sobre sus propiedades psicométricas en muestras españolas de pacientes de AP para así obtener datos más concluyentes que justifiquen de manera más sólida su uso para esos fines y en ese tipo de población.

Por tanto, como ya se adelantaba en el capítulo dedicado a los objetivos generales de esta tesis doctoral, el objetivo principal de este Estudio 2º fue profundizar en el análisis de las propiedades psicométricas del BDI-FS en muestras españolas de pacientes de AP.

Por un lado, se pretendía analizar algunas propiedades psicométricas importantes que no fueron examinadas en el Estudio 1º, en concreto, la estabilidad temporal o fiabilidad test-retest. La fiabilidad es una cuestión clave cuando se está adaptando o desarrollando un instrumento de medida, ya que es necesario cierto nivel de consistencia para poder evaluar la validez. La fiabilidad test-retest es uno de los indicadores más rigurosos de consistencia de una medida porque evalúa su estabilidad en el tiempo, pero su correcta estimación acarrea muchas dificultades, especialmente en el ámbito clínico. Una de las dificultades en este ámbito radica en que, con frecuencia, se observa que las puntuaciones de los tests psicopatológicos se atenúan en la segunda medida (Bumberry, Oliver y McClure, 1978; Hatzenbuehler, Parpal y Matthews, 1983; Oliver y Burkham, 1979; Piacentini et al., 1999), por lo que es

difícil desentrañar si los cambios que se producen entre la primera y segunda medición se deben a errores de medida del instrumento o, por el contrario, a cambios reales en la sintomatología del paciente debidos bien al curso natural de dichos síntomas o bien a los efectos terapéuticos del primer contacto clínico o a los efectos de la propia administración de tests (p. ej., aburrimiento, fatiga).

En este sentido, una de las cuestiones más importantes para evaluar adecuadamente la fiabilidad test-retest de un instrumento es seleccionar un intervalo para la segunda medida que se corresponda con la naturaleza del constructo que se está midiendo y con el marco temporal del instrumento. Por ejemplo, parecería obvio que en el caso de un instrumento que mida un rasgo de personalidad en adultos de 40 años, un constructo que se presupone relativamente estable a esa edad, se pueda evaluar la fiabilidad test-retest con intervalos amplios (p. ej., tres, seis o incluso 12 meses), mientras que en el caso de un instrumento como el BDI-FS que mide sintomatología depresiva, puesto que este constructo se presupone mucho más variable y cambiante, los intervalos para evaluar su fiabilidad test-retest debería necesariamente ser mucho más cortos (p. ej., un mes, dos semanas, una semana o incluso unos pocos días). En este último caso, además, si las instrucciones del instrumento preguntan al evaluado por su situación durante la última semana, parecería lógico que la fiabilidad test-retest de dicho instrumento se evaluara antes de que hubiera pasado una semana más, ya que si no se corre riesgo de que la sintomatología depresiva de la persona realmente haya cambiado y, por tanto, la segunda medida no esté realmente midiendo el mismo nivel de sintomatología que medía la primera medida; por el contrario, y con la misma lógica, si el instrumento pregunta por la situación del evaluado durante el último mes, entonces tendría sentido ampliar más días el intervalo de la segunda

medición, pero no más allá de que se hubiera acabado otro mes para que así ambos intervalos se solapen y, por tanto, se pueda evaluar la consistencia del instrumento para evaluar el mismo nivel de sintomatología depresiva.

Esta última cuestión ha sido a veces descuidada en la literatura empírica previa. Por ejemplo, Oliver y Burkham (1979) evaluaron la fiabilidad test-retest de la versión original del BDI, cuyas instrucciones preguntan por la situación actual del individuo durante el mismo día de su administración. Para ello, los investigadores utilizaron un intervalo para la segunda administración del inventario de 3 semanas. Por lo tanto, parece obvio que la sintomatología depresiva que presentaban los evaluados el mismo día de la segunda medición no tendría por qué corresponderse exactamente con la sintomatología que experimentaban el día en que se realizó la primera medida.

En este segundo estudio se trató de solventar este problema para que el mismo no confundiera la estimación de la fiabilidad test-retest del BDI-FS. Para ello, se realizó la segunda administración del BDI-FS a la semana, de forma que este intervalo se solapara con el marco temporal de su primera administración y, por tanto, hubiera cierta seguridad de que se estaba evaluando el mismo nivel de sintomatología depresiva. En este sentido, es importante recordar que las instrucciones del BDI-FS preguntan al individuo por su situación durante las dos últimas semanas, incluyendo el mismo día de su administración.

Es más, hasta ahora sólo se conoce de la existencia de un estudio publicado sobre la fiabilidad test-retest del BDI-FS, en concreto con una muestra de pacientes ancianos supervivientes a un infarto de miocardio (Healey, Kneebone, Carroll y Anderson, 2008), sin que tampoco se recogiera

información sobre dicha fiabilidad en el manual de la versión original del inventario (Beck et al., 2000). Por lo tanto, este segundo estudio de la presente tesis doctoral se constituye en la primera investigación sobre la fiabilidad test-retest del BDI-FS en una muestra de pacientes de AP.

Por otro lado, este segundo estudio trataba de ampliar la información sobre algunas propiedades psicométricas que ya fueron analizadas en el Estudio 1º, en particular, la validez convergente y la validez discriminante, pero extendiéndolas a otros instrumentos que no fueron empleados en el Estudio 1º y que también están validados en España y se usan con frecuencia en AP, al menos en otros países. Así, se comprobó la validez convergente del BDI-FS en relación con la Escala Autoaplicada para la Depresión de Zung (*Self-Rating Depression Scale* o SDS; Zung, 1965), adaptada en España por el equipo de investigación de Conde (Conde, Escriba e Izquierdo, 1970) y que constituye uno de los cuestionarios de referencia para la detección de la depresión en AP en varios países (Feightner y Worrall, 1990; Mulrow, Williams, Gerety, Ramirez, Montiel y Kerber, 1995), mientras que la validez discriminante del BDI-FS se examinó en relación con el Inventario de Ansiedad de Beck (*Beck Anxiety Inventory* o BAI; Beck, Epstein, Brown y Steer, 1988), adaptado en España por el equipo de investigación de Sanz (Sanz y Navarro, 2003; Magán, Sanz y García-Vera, 2008; Sanz, García-Vera y Fortún, 2011), sin duda uno de los instrumento de evaluación de la ansiedad más utilizado en la investigación psicológica (Piotrowski y Gallant, 2009) y que también ha sido utilizado con relativa frecuencia en AP (Beck, Steer, Ball, Ciervo y Kabat, 1997; Wetherell y Arean, 1997).

El Estudio 2º también trataba de replicar en una nueva muestra de pacientes de AP procedentes de otra ciudad española distinta a la ciudad en la que se hizo el Estudio 1º los resultados encontrados en este estudio respecto a las propiedades psicométricas de fiabilidad de consistencia interna y de validez factorial del BDI-FS, replicación que sentaría unas bases más sólidas para el empleo del inventario en AP en España.

Finalmente, en el Estudio 2º, tras combinar sus datos con los datos obtenidos en el Estudio 1º, se pretendía analizar las diferencias en las puntuaciones del BDI-FS en función de características demográficas básicas (sexo, edad, estado civil y nivel de estudios) y confeccionar unos baremos que tuvieran en cuenta esas diferencias y ayudaran a los clínicos a interpretar el nivel de sintomatología depresiva de los pacientes de AP.

7.2. Método

7.2.1. Participantes

En el presente estudio participaron 200 pacientes españoles de AP, tamaño muestral considerado como adecuado para la validación de instrumentos de evaluación psicológica según los estándares propuestos por Prieto y Muñiz (2000). Los participantes fueron una muestra incidental de los pacientes españoles que acuden a AP. En concreto, el estudio se llevó a cabo en los cuatro centros de atención primaria correspondientes a cada zona de salud de Palencia capital. La selección de los médicos de AP, dos por cada zona de salud, se realizó solicitando su colaboración voluntaria en turno vespertino. La derivación de los pacientes se propuso a través de un

procedimiento sistemático, seleccionándose el paciente correspondiente a cada 15 minutos de la lista de cada médico de AP. En el caso de que el paciente no accediese a participar o no cumpliese los criterios de inclusión, se escogió al paciente siguiente de la lista diaria de citas. A cada médico de AP se le asignó el mismo número de pacientes a derivar.

Se consideraron los siguientes criterios de inclusión: 1) estar adscrito al centro de atención primaria; 2) edad comprendida entre los 18 y los 80 años, y 3) no haber acudido a consulta por motivos exclusivamente administrativos ni por delegación de terceras personas.

Los criterios de exclusión que se utilizaron fueron los siguientes (en función del criterio clínico del médico de AP): 1) tener patología psicótica, demencia o enfermedad concurrente que hiciese imposible la aplicación de las pruebas del estudio; 2) no tener capacidades cognitivas básicas de atención o concentración, y 3) no tener una comprensión suficiente del castellano.

Para el presente estudio se seleccionaron 200 pacientes que cumplieran los criterios de inclusión y que no cumplieran los criterios de exclusión anteriormente mencionados. Algunos pacientes no contestaron todas las preguntas sociodemográficas ni cumplimentaron todas las preguntas de los cuestionarios, por lo que el número final de participantes para cada análisis varió entre los 198 participantes que completaron totalmente el BDI-FS y los 188 que completaron totalmente la SDS. En cada análisis y en cada tabla de resultados se indicará el número concreto de participantes en que están basados.

La muestra de 200 pacientes de atención primaria estaba compuesta, casi a partes iguales, por mujeres (49,5%) y varones (48,5%), y su rango de

edad oscilaba entre 18 y 80 años (edad media = 44,11 años; $DT = 20,28$ años). Atendiendo a las características demográficas más frecuentes en la muestra, su perfil sociodemográfico se correspondía con el de un paciente con una edad en la franja de 30-44 años (27,5%), casado o conviviendo de manera estable (59,5%), con estudios primarios (36,0%), trabajando (56%) y con profesiones muy diversas. En la Tabla 7.1 se presentan con mayor detalle las principales características demográficas de la muestra de pacientes de AP del presente estudio.

Tabla 7.1. Características demográficas de la muestra de pacientes de atención primaria del Estudio 2º

<i>Característica</i>	<i>Frecuencia</i>	<i>Porcentaje (%)</i>
Sexo:		
- Mujeres	99	49,5
- Varones	97	48,5
- No contesta	4	2,0
Edad:		
- 18-29 años	31	15,5
- 30-44 años	55	27,5
- 45-59 años	54	27,0
- 60 años o más	42	21,0
- No contesta	18	9
Estado civil:		
- Soltero	55	27,5
- Casado o conviviendo con pareja de forma estable	119	59,5
- Separado o divorciado	13	6,5
- Viudo	9	4,5
- No contesta	4	2

Característica	Frecuencia	Porcentaje (%)
Formación académica:		
- No tiene estudios	12	6
- Estudios primarios, EGB o equivalente	72	36,0
- Bachillerato, FP o equivalente	68	34,0
- Estudios universitarios	43	21,5
- No contesta	5	2,5
Situación laboral:		
- Trabaja	112	56
- Jubilado o pensionista	28	14
- Parado	9	4,50
- Estudiante	20	10
- Sus labores	20	10
- Incapacidad laboral transitoria	1	0,5
- Otra situación	5	2,5
- No contesta	5	2,5

Tabla 7.1. Características demográficas de la muestra de pacientes de atención primaria del Estudio 2º (continuación)

Tabla 7.1. Características demográficas de la muestra de pacientes de atención primaria del Estudio 2º (Continuación)

Característica	Frecuencia	Porcentaje (%)
Profesión:		
- Profesionales o técnicos	25	12,5
- Altos directivos de la administración pública o empresa privada	2	1,0
- Propietarios o gerentes de hostelería, comercio o agricultura	8	4,0
- Mandos intermedios, profesores, técnicos de la administración	21	10,5
- Capataces y personal encargado en empresas o actividades de servicio	3	1,5
- Personal administrativo	22	11,0
- Comerciantes, vendedores y similares	3	1,5
- Personal de servicio	24	12,0
- Trabajadores cualificados o semicualificados	21	10,5
- Fuerzas armadas	5	2,5
- Estudiante	21	10,5
- Sus labores	22	11,0
- Otra ocupación	18	9,0
- No contesta	5	2,5

7.2.2. Instrumentos

- **Inventario de Depresión de Beck–Cribado Rápido** (BDI-FS; Beck et al., 2000). El instrumento objeto de esta tesis ya ha sido descrito con detalle en el Capítulo 4, tanto en su versión original como en la versión española que se está validando.

- **Inventario de Ansiedad de Beck (BAI)** o *Beck Anxiety Inventory* (Beck, Epstein, Brown y Steer, 1988). Inventario cuyo objetivo es valorar la presencia

de síntomas de ansiedad y su gravedad. También se ha desarrollado con el fin de discriminar de manera más fiable entre ansiedad y depresión.

Se trata de un inventario autoaplicado formado por 21 ítems. Cada ítem del BAI recoge un síntoma de ansiedad y para cada uno de ellos la persona evaluada debe indicar el grado en que se ha visto afectado por el mismo durante la última semana utilizando para ello una escala tipo Likert de cuatro puntos: 0 (Nada en absoluto), 1 (Levemente, no me molestó mucho), 2 (Moderadamente, fue muy desagradable pero podía soportarlo), ó 3 (Gravemente, casi no podía soportarlo). El tiempo de administración aproximado es de 5 minutos.

Respecto a su corrección e interpretación, cada ítem se valora de 0 a 3 puntos en función de la respuesta dada por la persona evaluada y, tras sumar directamente la puntuación de cada ítem, se obtiene una puntuación total cuyo rango es de 0 a 63. En la segunda edición del manual original del BAI, Beck y Steer (1993) proponen puntos de corte algo más bajos respecto a la primera edición para definir las distintas categorías de gravedad de la ansiedad: 0-7 = “normal”; 8-15 = “leve”; 16-25 = “moderada”, y 26-63 = “grave”.

Existen dos adaptaciones en español diseñadas para los hispanos de Estados Unidos de América, una creada por la empresa editora del BAI original (*Psychological Corporation*, 1993) y otra por Novy, Stanley, Averill y Daza (2001). También existe una versión española adaptada a la población mexicana (Robles, Varela, Jurado y Páez, 2001). En España Sanz y Navarro (2003) realizaron una nueva versión española del BAI y analizaron sus propiedades psicométricas en una muestra española de estudiantes universitarios, mientras que Magán, Sanz y García-Vera (2008) lo hicieron en

una muestra de adultos de la población general española y Sanz, García-Vera y Fortún (2011) en una muestra de pacientes psicopatológicos. Por tanto, esta es la única versión que puede considerarse, por ahora, una versión española del BAI adaptada para la población de España.

En relación con sus propiedades psicométricas, el BAI se construyó con la intención de disponer de una medida de ansiedad clínica que a su vez permitiera discriminar la ansiedad de la depresión. Según el análisis de su validez de contenido realizado por Sanz y Navarro (2003), el BAI se distingue, en cuanto a su contenido, por evaluar sobre todo síntomas fisiológicos: 14 de sus 21 ítems (el 67%) se refieren a síntomas fisiológicos, mientras que sólo 4 de sus ítems evalúan aspectos cognitivos y 3 aspectos afectivos. Por otro lado, 19 de los 21 ítems del BAI (el 90%) se refieren a síntomas característicos de las crisis de angustia o pánico. Teniendo en cuenta un concepto de ansiedad clínica tan consensuado como el que aparece reflejado en el DSM-IV, el BAI cubre 13 de los 37 síntomas distintos que definen los trastornos de ansiedad primarios en el DSM-IV. Sin embargo, dado que 8 de esos 37 síntomas también forman parte de los criterios sintomáticos que definen los trastornos depresivos en el DSM-IV, cabe concluir que los ítems del BAI reflejan el 45% de los síntomas específicos de la ansiedad clínica según el DSM-IV.

Fiabilidad: La fiabilidad medida a través de su consistencia interna es adecuada a excelente. Los autores del BAI encontraron un alfa de Cronbach igual a 0,92 en pacientes psicopatológicos. Posteriormente De Ayala, Vonderharr Carlson y Kim (2005) en su metanálisis de 52 estudios hallaron un coeficiente alfa medio de 0,92 con muestras de pacientes con trastornos psicopatológicos, de 0,88 con muestras de la población general y de 0,89 con

muestras de estudiantes universitarios. En España, la adaptación española de Sanz y Navarro (2003) ha obtenido un coeficiente alfa de 0,88 en una muestra de 590 estudiantes universitarios españoles, de 0,93 en una muestra de 249 adultos de la población general española (Magán et al., 2008) y de 0,90 en una muestra de 307 pacientes con trastornos mentales (Sanz et al., 2011). La *fiabilidad test-retest* al cabo de una semana que se obtuvo en el estudio original con una muestra de 83 pacientes ambulatorios con trastornos psicopatológicos fue de 0,75, el cual se reduce en el meta-análisis de De Ayala et al. (2005) a un 0,66, aunque con un incremento en el intervalo de tiempo medio (32 días).

Validez. En relación con la validez convergente, los autores originales encontraron que el BAI correlacionó 0,51 con la Escala de Valoración de la Ansiedad de Hamilton en una muestra de pacientes psicopatológicos, mientras que en el estudio de Creamer, Foran y Bell (1995) con estudiantes universitarios, el BAI mostró correlaciones de 0,57-0,68 y de 0,56-0,64 con las escalas de Estado y Rasgo, respectivamente, del Cuestionario de Ansiedad Estado-Rasgo (STAI) de Spielberger. Estos datos hacen pensar en una adecuada validez convergente del BAI con otras medidas de ansiedad.

En cuanto a la *validez discriminante* respecto a la depresión, aunque la literatura científica muestra también correlaciones moderadas entre el BAI y diferentes medidas de depresión como ocurre, por ejemplo, en los estudios de la adaptación española en los que el BAI correlacionaba 0,63, 0,58 y 0,56 con el BDI-II en población general, estudiantes universitarios y pacientes psicopatológicos, respectivamente, sin embargo, los análisis factoriales de los ítems del BAI y del BDI-II revelaban en tales estudios dos factores claramente diferenciados de ansiedad y depresión en los que saturaban los ítems

respectivos y, por tanto, sugieren que las correlaciones entre los dos instrumentos pueden deberse más a la relación entre los constructos de ansiedad y depresión que a un problema de falta de validez discriminante por parte del BAI (Sanz y Navarro, 2003; Magán et al., 2008; Sanz et al., 2011).

En la muestra de pacientes de AP del presente estudio, el BAI obtuvo un índice de fiabilidad de consistencia interna (*alfa* de Cronbach) igual a 0,90, valor excelente según los estándares propuestos por Prieto y Muñiz (2000) y similar a los obtenidos en estudios españoles previos con muestras de estudiantes universitarios, de la población general o de pacientes con trastornos mentales (Sanz y Navarro, 2003; Sanz et al., 2011; Magán et al., 2008).

- **Escala Autoaplicada para la Depresión de Zung (SDS)** o *Self-Rating Depression Scale* (Zung, 1965). La SDS es un cuestionario autoaplicado para la detección y la medida de la gravedad de la depresión que consta de 20 ítems que exploran síntomas depresivos y hacen referencia a la frecuencia de la presentación de los síntomas: “raramente”, “a veces”, “frecuentemente” o “siempre”. De acuerdo a sus instrucciones, el marco de referencia temporal para dicha evaluación es el momento actual. En este estudio se utilizó la adaptación española de la SDS realizada por el equipo de investigación de Conde (Conde, Escribá e Izquierdo, 1970).

Corrección e interpretación: La SDS proporciona una puntuación total de gravedad de la depresión que es la suma de los 20 ítems, teniendo en cuenta que cada ítem se puntúa de 1 a 4 puntos. Para evitar el sesgo de aquiescencia, la SDS alterna ítems redactados en sentido directo, indicando un mayor nivel de depresión, con ítems redactados en sentido inverso, indicando un menor

nivel de depresión. Los ítems directos, con contenido, pues, negativo, son los ítems 1, 3, 4, 7, 8, 9, 10, 13, 15 y 19. Los ítems inversos, con contenido, pues, positivo, son los ítems 2, 5, 6, 11, 12, 14, 16, 17, 18 y 20. Los puntos de corte para la versión española son: 20-35 = “depresión ausente”; 36-51 = “depresión subclínica y variantes normales”; 52-67 = “depresión media-grave”, y 68-80 = “depresión grave” (Conde et al., 1974).

Las propiedades psicométricas de la adaptación española del SDS son adecuadas y parecidas a las de la versión original en muestras muy diversas (Conde et al., 1970; Conde y Esteban, 1975). Por ejemplo, en cuanto a la fiabilidad de consistencia interna, Aragonés et al. (2001) encontraron en una muestra de 350 pacientes de AP un coeficiente *alfa* de Cronbach igual a 0,78, mientras que cuando dicha fiabilidad fue estimada a través de la correlación entre dos mitades $r_{x,x'}$ (corregido), obtuvieron un coeficiente igual a 0,80.

En cuanto a la validez, en el estudio anterior de Aragonés et al. (2001) también se analizó la validez de criterio de la versión española de la SDS, para lo cual se corrigió la SDS transformando su puntuación bruta en un índice porcentual (índice SDS = puntuación bruta x 100/80) y se consideraron como puntos de corte para distintos niveles de gravedad de la depresión los habituales cuando se utiliza el índice porcentual de la SDS, es decir, 50% para la depresión leve, 60% para la depresión moderada, y 70% para la depresión grave. Aragonés et al. (2001) encontraron que, en comparación al diagnóstico de trastorno depresivo obtenido con la SCID-I, la puntuación igual o mayor a 54% obtuvo una sensibilidad de 0,85 y una especificidad de 0,86 para detectar trastornos depresivos, cifras que les llevaron a considerar a la SDS un cuestionario eficaz y con unas propiedades psicométricas aceptables para la detección de la depresión en AP en España.

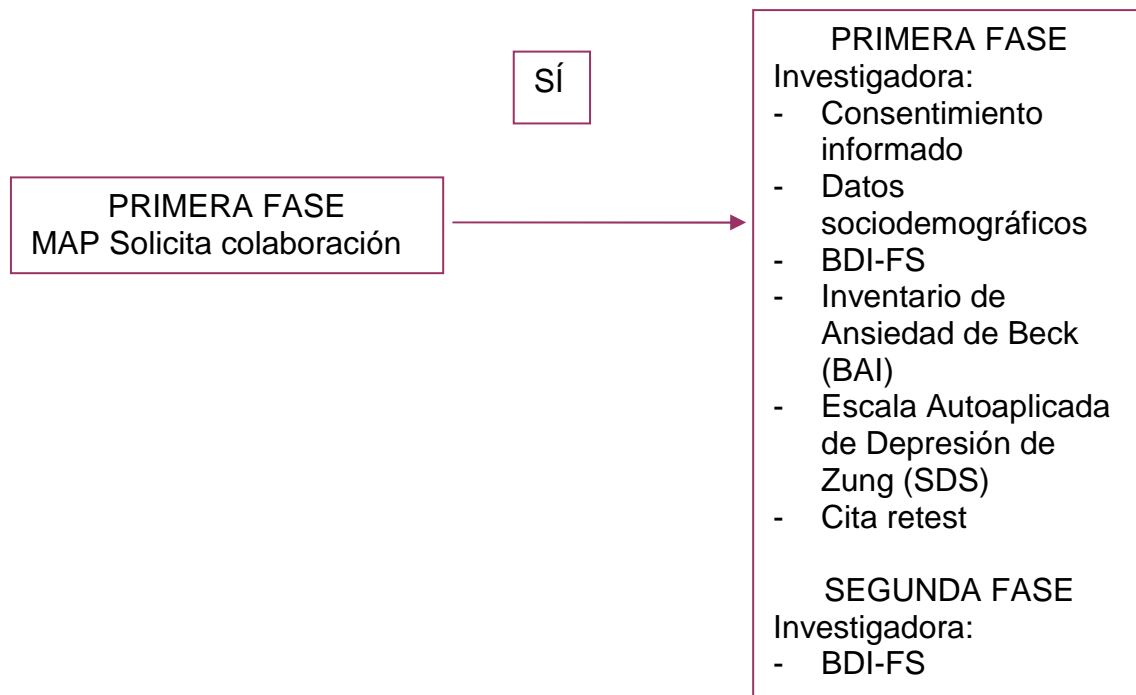
En el presente estudio la fiabilidad de consistencia interna de la SDS, en términos del coeficiente *alfa* de Cronbach, fue igual a 0,78, lo que representa un índice adecuado según los estándares propuestos por Prieto y Muñiz (2000) y similar al obtenido en el estudio previo de Aragonés et al. (2001) con pacientes españoles de AP.

7.2.3. Procedimiento

El estudio se realizó en dos fases, que se resumen en la Figura 7.1.

1) Fase 1. En esta fase se evaluaron 200 pacientes que fueron elegidos de forma sistemática en las consultas de AP mediante la selección de las citas que correspondían a las horas en punto y a las citas de las medias horas. A estos pacientes, el médico de AP les expuso la existencia de una investigación para la exploración de alteraciones emocionales y les solicitó su colaboración. La identificación de la paciente se realizó con una clave formada por el número de paciente enviado de forma correlativa y el número de identificación asignado al médico de AP que lo derivaba. Posteriormente el médico de AP derivó al paciente al despacho de la investigadora. Ésta le explicó al paciente la naturaleza del estudio, obtuvo el consentimiento informado del paciente para su participación en este estudio y le explicó que podían rechazar proseguir con la investigación en cualquier momento de la misma. Posteriormente, la investigadora facilitó al paciente un protocolo formado por el BDI-FS, el BAI, la SDS y un cuestionario de variables sociodemográficas.

Figura 7.1. Resumen del procedimiento del segundo estudio



2) Fase 2. Una semana posterior a la primera cita, a los 200 pacientes se les administró vía telefónica o presencial (según preferencia de la persona objeto de la investigación) el BDI-FS para obtener el retest de la prueba.

Los datos fueron recogidos por dos investigadoras licenciadas en Psicología, una de ellas (la autora de la presente tesis doctoral) especialista en Psicología Clínica y la segunda residente de Psicología Clínica de segundo año. Ambas investigadoras tuvieron el mismo número de personas derivadas por día. Las dos investigadoras habían recibido entrenamiento específico en el uso del BDI-FS, del BAI y de la SDS, acordando usar el mismo procedimiento en las instrucciones administradas durante el protocolo.

7.2.4. Diseño y análisis de los datos

Este estudio respondió a un diseño correlacional transversal de tipo instrumental. Para el análisis de los datos se utilizó el paquete estadístico *SPSS 15.0 para Windows XP*. Con este paquete estadístico, se analizaron de forma descriptiva las distribuciones de las puntuaciones calculándose la media, la desviación típica, los índices de curtosis y simetría, y el test de normalidad de Kolmogorov-Smirnov.

La fiabilidad de consistencia interna se midió a través del coeficiente *alfa* de Cronbach, mientras que la fiabilidad test-retest fue calculada mediante el coeficiente de correlación de Pearson entre las puntuaciones del BDI-FS de las dos administraciones del inventario separadas por un lapso de tiempo de una semana. Se analizó también si existían diferencias estadísticamente significativas entre las medias de ambas aplicaciones mediante una prueba *t* de Student para medidas repetidas.

La validez convergente y discriminante del BDI-FS se analizó calculando sus coeficientes de correlación de Pearson con la SDS y el BAI, y calculando las diferencias entre dichos coeficientes de correlación mediante la *Z* de Steiger. Para analizar la validez discriminante del BDI-FS también se realizaron varios análisis factoriales con las puntuaciones de los ítems de los tres instrumentos (BDI-FS, SDS y BAI) según el método de extracción de factores de ejes principales y de rotación oblicua de factores *promax*.

El mismo método de análisis factorial, pero solamente sobre los ítems del BDI-FS, se utilizó para evaluar la validez factorial del inventario.

Para comprobar si existían diferencias en las puntuaciones del BDI-FS en función de las características demográficas de los pacientes de AP (sexo,

edad, estado civil y niveles de estudios), se analizaron conjuntamente los datos del Estudio 1º y los del presente estudio, y, en el caso del estado civil, se unió en una sola categoría los pacientes divorciados o separados y los viudos, ya que estos últimos suponían menos de 20 casos. Se realizaron varios ANCOVA sobre la puntuación total del BDI-FS tomando como factor intersujeto, sucesivamente, el sexo, el grupo de edad (cuatro grupos o niveles: 16-29 años, 30-44 años, 45-59 años, y 60 o más años), el estado civil (tres grupos o niveles: soltero, casado o conviviendo con pareja de manera estable, y separado o divorciado o viudo) y el nivel de estudios (cuatro grupos o niveles: sin estudios, estudios primarios o EGB o equivalente, bachillerato o FP o equivalente, y estudios universitarios). En cada uno de esos ANCOVA, se incluyó como covariables el resto de variables demográficas que no constituirían el factor intersujeto. Por ejemplo, para el ANCOVA del factor intersujetos sexo, se incluyó como covariables la edad, el nivel de estudios y el estado civil, y así sucesivamente para el resto de factores intersujetos. En todos esos ANCOVA, cuando la edad fue introducida como covariable se hizo como variable continua, no como variable de grupos de edad, y, en el caso del estado civil, al ser una variable categórica con tres categorías, se crearon dos variables estadísticas, ficticias o de diseño (*dummy variables*) que fueron las que se incluyeron como covariables: separado/divorciado/viudo frente a no separado/no divorciado/no viudo, y casado/conviviendo establemente con pareja frente a no casado/no conviviendo establemente con pareja.

En los ANCOVA anteriores, para estimar el tamaño de los efectos del factor intersujeto se calculó el estadístico η^2 o eta cuadrado. Levine y Hullet (2002) argumentan que este estadístico es preferible η^2 parcial, el estadístico que el programa SPSS calcula cuando realiza un ANOVA o un ANCOVA. El

cálculo de η^2 se llevó a cabo a partir de la sumas de cuadrados del ANCOVA que ofrece el SPSS. Según Cohen (1988, p. 283), un valor de η^2 igual a 0,0099 representa un tamaño del efecto pequeño, 0,0588 un tamaño del efecto medio, y 0,1379 un tamaño del efecto grande.

Cuando los ANCOVA incluían como factores intersujetos variables sociodemográficas con más de dos grupos (grupo de edad, estado civil y nivel de estudios) y fueron estadísticamente significativos, se realizaron pruebas *t* de Student de comparación entre pares de grupos con los niveles de significación corregidos por el método de Bonferroni, y el tamaño del efecto de las diferencias entre grupos se estimó mediante el estadístico *d* de Cohen (o diferencia estandarizada en desviaciones típicas). Según Cohen (1988), un valor de *d* igual a 0,20 representa un tamaño del efecto pequeño, 0,50 un tamaño del efecto medio, y 0,80 un tamaño del efecto grande.

Finalmente, para la construcción de baremos se calcularon los percentiles más frecuentes (1, 5, 10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 95 y 99) para la muestra conjunta del Estudio 1º y del presente estudio, así como para dicha muestra en función del sexo.

7.3. Resultados

7.3.1. Distribución de las puntuaciones del BDI-FS

En la muestra de pacientes de AP de este segundo estudio ($N = 198$), se encontró una puntuación total media en el BDI-FS de 3,02, con una desviación típica de 3,16, una puntuación máxima de 17 y una mínima de 0 puntos.

Al analizar cómo se distribuían las puntuaciones del BDI-FS mediante el test de normalidad de Kolmogorov-Smirnov, este se reveló como estadísticamente significativo ($p < 0,001$), lo que implicaba que las puntuaciones del BDI-FS no se distribuían de forma normal. En esta línea, los índices de asimetría y curtosis, 1,82 y 4,28, respectivamente, también aparecieron fuera de los rangos de normalidad (± 1). Considerando el primero, se encontró una asimetría positiva, es decir, los datos fundamentalmente se aglutinaban a la izquierda de la media. En cuanto a la curtosis, se encontró una elevada concentración de datos alrededor de la media, lo que se aproximaba a una distribución leptocúrtica. En la Figura 7.2 se muestra gráficamente en qué medida la distribución de puntuaciones del BDI-FS no se ajustaba a una distribución normal.

Figura 7.2. Distribución de las puntuaciones total del BDI-FS en la muestra de pacientes de atención primaria del Estudio 2º

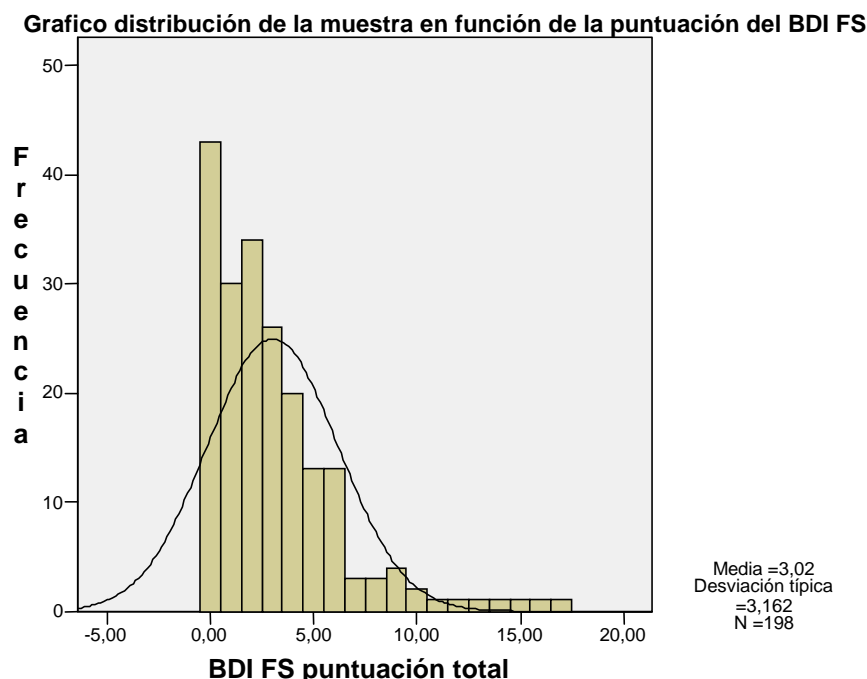


Tabla 7.2. Distribución de la muestra de pacientes de atención primaria del Estudio 2º en función de los niveles de gravedad del BDI-FS

Rangos cuantitativos de depresión	Frecuencia	% de datos válidos
Depresión mínima o ausente (BDI-FS = 0-3)	133	67,2
Depresión leve (BDI-FS = 4-6)	46	23,2
Depresión moderada (BDI-FS = 7-9)	10	5,1
Depresión grave (BDI-FS = 10-21)	9	4,5
Total de datos válidos	198	100,0
Datos perdidos	2	
Total	200	

Tabla 7.3. Características psicométricas de los ítems del BDI-FS en la muestra de pacientes de atención primaria del Estudio 2º (N = 198)

Ítem del BDI-FS	Gravedad del síntoma -				Media	DT	$r_{\text{ítem-total}}$
	% de respuesta						
	0	1	2	3			
1. Tristeza	75,8	19,7	3	1,5	0,3	0,6	0,54
2. Pesimismo	61,3	26,1	7,0	5,5	0,5	0,8	0,61
3. Sentimientos de fracaso	74,4	22,1	2,0	1,5	0,3	0,5	0,63
4. Pérdida de placer	51,3	43,2	3,0	2,5	0,5	0,6	0,58
5. Insatisfacción con uno mismo	68,3	24,1	5	2,5	0,4	0,7	0,54
6. Autocríticas	43,7	44,2	7,5	4,5	0,7	0,7	0,51
7. Pensamientos de suicidio	87,9	12,1	0	0	0,1	0,3	0,38

Teniendo en cuenta los criterios de gravedad de sintomatología depresiva propuestos por los autores del BDI-FS (Beck et al., 2000), la distribución de las puntuaciones del BDI-FS en la presente muestra de pacientes de AP indicaba que la mayoría de dichos pacientes mostraban

niveles de depresión ausente o leve, siendo más infrecuente la sintomatología depresiva moderada o grave (véase la Tabla 7.2).

Al realizar un análisis por ítems del BDI-FS, la mayoría de los pacientes puntuaron en los niveles de gravedad más bajos para todos y cada uno de dichos ítems, siendo el ítem que alcanzó una mayor puntuación media el de Autocríticas y el que obtuvo una menor puntuación media el de Pensamientos de suicidio (véase la Tabla 7.3).

7.3.2. Fiabilidad

Un índice fundamental para ver la calidad de un test es su fiabilidad. En este segundo estudio se han obtenido dos índices de fiabilidad, el de consistencia interna y el de estabilidad temporal o test-retest.

7.3.2.1. Fiabilidad de consistencia interna

Para la muestra de pacientes de AP del presente estudio se obtuvo un *alfa* de Cronbach igual a 0,80, valor considerado como bueno según los estándares de Prieto y Muñiz (2000). Por otro lado, la media de las correlaciones entre ítems fue 0,37 con un mínimo de 0,21 y un máximo de 0,59, valores todos ellos adecuados según los estándares propuestos por Briggs y Cheek (1986; véase también Bollen y Lennox, 1991).

Considerando como índice de consistencia interna de los ítems su correlación ítem-total corregida (véase la Tabla 7.3), es decir, la correlación entre el ítem y la puntuación total sin el concurso de ese ítem, todos los ítems del BDI-FS mostraron buenos índices ya que en todos los casos se superó el

valor estándar de 0,30 (Nunnally y Bernstein, 1995) o de 0,35 (Cohen y Manion, 1990), siendo el valor máximo el del ítem de Fracaso (0,63) y el menor el del ítem de Suicidio (0,38).

7.3.2.2. Fiabilidad test-retest

En este estudio se pretendía examinar la fiabilidad test-retest del BDI-FS a la semana y, finalmente, el intervalo real en que se realizó el retest osciló entre 6 días y 9 días, con una media de 7,12 días y una desviación típica de 0,37. De los 200 pacientes evaluados inicialmente, se consiguió contactar de nuevo con 167 pacientes y obtener datos válidos del test-retest para 166 pacientes. Esta última cifra supone una pérdida de datos del 17% que fue debida a muy diferentes causas: no acudir a la cita programada para la semana posterior, no aceptar la llamada telefónica, rehusar responder de nuevo por teléfono o presencialmente, encontrarse el número de teléfono proporcionado fuera de servicio o ser falso, no contestar a todos los ítems, etc.

La correlación de Pearson entre las puntuaciones totales del BDI-FS de la primera aplicación y la realizada una semana después fue igual a 0,89. Este coeficiente fue estadísticamente significativo ($p < 0,001$) y representa un índice de fiabilidad test-retest excelente según los estándares de Prieto y Muñiz (2000), por lo que indicaría una muy buena estabilidad de las puntuaciones del BDI-FS a lo largo de una semana.

Como se comentaba en el epígrafe introductorio de este capítulo, es habitual encontrar un descenso en las puntuaciones de los instrumentos psicopatológicos cuando se administran por segunda vez y, efectivamente, este fue el caso en el presente estudio. Una prueba t de Student para medidas

repetidas reveló que la media del BDI-FS en el retest fue significativamente menor que la media obtenida durante la primera aplicación [2,83 frente a 3,06; $t(165) = 2,03$, $p = 0,044$], aunque la diferencia entre ambas media fue muy pequeña en términos del tamaño del efecto (d de Cohen = 0,07) y no parece clínicamente relevante.

7.3.3. Validez convergente

El BDI-FS correlacionó de forma positiva y estadísticamente significativa ($p < 0,001$) con la Escala Autoaplicada para la Depresión de Zung (SDS) obteniendo un coeficiente de correlación de Pearson igual a 0,71, lo que representa un índice de validez convergente excelente según los estándares de Prieto y Muñiz (2000).

7.3.4. Validez discriminante

En la presente muestra de pacientes de AP se encontró una correlación positiva y significativa entre el BDI-FS y el BAI igual a 0,57 ($p < 0,001$). No obstante, y como cabría de esperar, la correlación que mostró el BDI-FS con el BAI fue menor que la correlación que mostró con la SDS ($r = 0,57$ frente a 0,71), y la diferencia fue estadísticamente significativa [Z de Steiger ($N = 181$) = 3,02, $p < 0,003$].

Es más, al calcular la correlación de la otra medida de depresión utilizada en la presente muestra, la SDS, con el BAI se obtuvo un coeficiente igual a 0,66 ($p < 0,001$), y comparando este coeficiente de correlación con el obtenido entre el BDI-FS y el BAI, se podría inferir que en esta muestra de

pacientes de atención primaria el BDI-FS, en comparación a la SDS, correlacionó menos con el BAI y, por ende, discriminó mejor respecto a la ansiedad medida por el BAI ($r = 0,57$ frente $0,66$), ya que de hecho la diferencia entre ambos coeficientes de correlación fue estadísticamente significativa [Z de Steiger ($N = 181$) = $2,25$, $p < 0,024$].

A pesar de estos indicios de buena validez convergente-discriminante por parte del BDI-FS, es cierto que la existencia de un coeficiente de correlación discriminante respecto al BAI superior a $0,50$, valor que se considera un efecto de tamaño grande según los estándares para las correlaciones de Cohen (1988), podría, en principio, sugerir falta de validez discriminante por parte del BDI-FS. No obstante, dado el solapamiento de las definiciones de ansiedad y depresión (Sanz y Navarro, 2003), cierta correlación entre sus medidas es esperable. Por tanto, la cuestión debería plantearse en términos de si el BDI-FS, a pesar de su alta correlación con medidas de ansiedad como el BAI, permite medir una sintomatología afectiva diferente de la ansiedad.

Para responder a esa pregunta se realizó un análisis factorial con todos los ítems del BDI-FS, de la SDS y del BAI, ya que además tanto el test de Kaiser-Meyer-Olkin de adecuación muestral ($KMO = 0,86$) como la prueba de esfericidad de Bartlett ($\chi^2 = 4064,02$, $p < 0,001$) indicaban que el modelo factorial era adecuado para esos datos. Puesto que se trataba de examinar si el BDI-FS mide síntomas afectivos diferentes de la ansiedad, en el análisis factorial se extrajeron dos factores mediante ejes principales y se rotaron con el método oblicuo *promax*. Como puede verse en la Tabla 7.4, la matriz de configuración resultante reflejó claramente la existencia de un factor

de ansiedad y otro de depresión, los cuales correlacionaban entre sí de forma alta ($r = 0,65$), pero quedaban claramente definidos por los ítems de sus correspondientes instrumentos.

Efectivamente, considerando como despreciables las saturaciones menores de 0,25 en la matriz de configuración y como definitorias las saturaciones iguales o mayores de 0,40 en dicha matriz, todos los ítems que definían el primer factor (ansiedad) pertenecían al BAI (cuyos ítems saturaban entre 0,27 y 0,83), con la única excepción del ítem “palpitaciones” de la SDS que mostraba una saturación mayor de 0,40 en ese primer factor (excepción que parece lógica dado que la presencia de palpitaciones o taquicardias es un síntoma más frecuente del constructo de ansiedad que del de depresión). Es más, 19 de los 21 ítems del BAI presentaban saturaciones mayores de 0,40 en ese primer factor de ansiedad (véase la Tabla 7.4). De forma similar, todos los ítems que definían el segundo factor (depresión) pertenecían al BDI-FS (cuyos ítems saturaban entre 0,85 y 0,27) o a la SDS (cuyos ítems saturaban entre -0,07 y 0,70), y 6 de los 7 ítems del BDI-FS presentaban saturaciones mayores de 0,40 en ese segundo factor (véase la Tabla 7.4). Contrariamente a lo que cabría esperar, los índices de validez de la SDS en este análisis factorial fueron peores que los del BDI-FS, ya que tan sólo 9 de sus 20 ítems mostraron saturaciones mayores de 0,40 en ese segundo factor. No obstante, es importante subrayar, que salvo el ítem “palpitaciones”, ninguno de los ítems de la SDS saturaban de forma significativa ($\geq 0,40$) en el primer factor de ansiedad (véase la Tabla 7.4).

Tabla 7.4. Análisis factorial conjunto de los ítems del BDI-FS, SDS y BAI: matriz de configuración de la solución bifactorial

Ítem	Instrumento	Factor 1	Factor 2
1. Tristeza	BDI-FS	–	0,536
2. Pesimismo	BDI-FS	-0,273	0,855
3. Sentimientos de fracaso	BDI-FS	–	0,686
4. Pérdida de placer	BDI-FS	–	0,619
5. Insatisfacción con uno mismo	BDI-FS	–	0,595
6. Auto-críticas	BDI-FS	–	0,473
7. Pensamientos o deseos de suicidio	BDI-FS	0,260	0,268
1. Tristeza	SDS	0,261	0,563
2. Mejoría matutina	SDS	–	–
3. Llanto	SDS	–	0,435
4. Problemas de sueño	SDS	0,266	–
5. Sin cambios de apetito	SDS	–	0,413
6. Sin cambios en interés por el sexo	SDS	–	0,376
7. Pérdida de peso	SDS	–	–
8. Estreñimiento	SDS	–	0,322
9. Palpitaciones	SDS	0,521	–
10. Cansancio	SDS	0,254	0,457
11. Cabeza despejada	SDS	–	–
12. Hacer las cosas con misma facilidad	SDS	–	0,534
13. Agitación e intranquilidad	SDS	–	–
14. Esperanza y confianza en el futuro	SDS	–	0,698
15. Irritabilidad	SDS	–	0,295
16. Facilidad para tomar decisiones	SDS	–	0,382
17. Sentimientos de ser útil	SDS	–	0,438
18. Sentimientos de vida plena- agradable	SDS	–	0,472

Nota. *N* = 181. No se muestran las saturaciones < 0,25 y las saturaciones > 0,40 se muestran en negrita.

Tabla 7.4. Análisis factorial conjunto de los ítems del BDI-FS, SDS y BAI: matriz de configuración de la solución bifactorial (Continuación)

Ítem	Instrumento	Factor 1	Factor 2
19. Ideación suicida	SDS	–	0,309
20. Placer y disfrute con las mismas cosas	SDS	–	0,550
1. Hormigueo o entumecimiento	BAI	0,490	–
2. Sensación de calor	BAI	0,508	–
3. Debilidad en las piernas	BAI	0,315	0,345
4. Incapacidad para relajarme	BAI	0,593	–
5. Miedo a que suceda lo peor	BAI	0,527	–
6. Mareos o vértigos	BAI	0,522	–
7. Palpitaciones o taquicardia	BAI	0,476	–
8. Sensación de inestabilidad	BAI	0,641	–
9. Sensación de estar aterrorizado	BAI	0,275	–
10. Nerviosismo	BAI	0,425	0,298
11. Sensación de ahogo	BAI	0,682	–
12. Temblor de manos	BAI	0,664	–
13. Temblor generalizado	BAI	0,823	–
14. Miedo a perder el control	BAI	0,581	–
15. Dificultad para respirar	BAI	0,677	–
16. Miedo a morir	BAI	0,459	–
17. Estar asustado	BAI	0,697	–
18. Indigestión o molestia abdominal	BAI	0,490	–
19. Sensación de desmayarse	BAI	0,827	-0,332
20. Rubor facial	BAI	0,701	–
21. Sudoración	BAI	0,631	–

Nota. $N = 181$. No se muestran las saturaciones $< 0,25$ y las saturaciones $> 0,40$ se muestran en negrita.

Para confirmar estos resultados que sugerían una buena validez discriminante del BDI-FS con respecto al BAI, se realizaron dos análisis factoriales más. En primer lugar, puesto que en el análisis factorial anterior el test o gráfico de sedimentación (*scree test*) de Cattell recomendaba una

solución de 3 factores (véase la Figura 7.3), se extrajeron tres factores mediante ejes principales y se rotaron con el método oblicuo *promax*. De hecho, el test de mínima correlación parcial promediada de Velicer (MAP), realizado mediante el programa de instrucciones de SPSS de O'Connor (2000), indicaba también una solución trifactorial para el análisis factorial conjunto de los ítems de los tres instrumentos (BDI-FS, SDS y BAI). Los estudios de simulación muestran que el MAP es uno de los mejores métodos para evaluar la dimensionalidad de una matriz de datos (Zwick y Velicer, 1986).

Figura 7.3. Gráfico de sedimentación de Cattell de los factores obtenidos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS, la SDS y el BAI de la muestra de pacientes de atención primaria del Estudio 2º

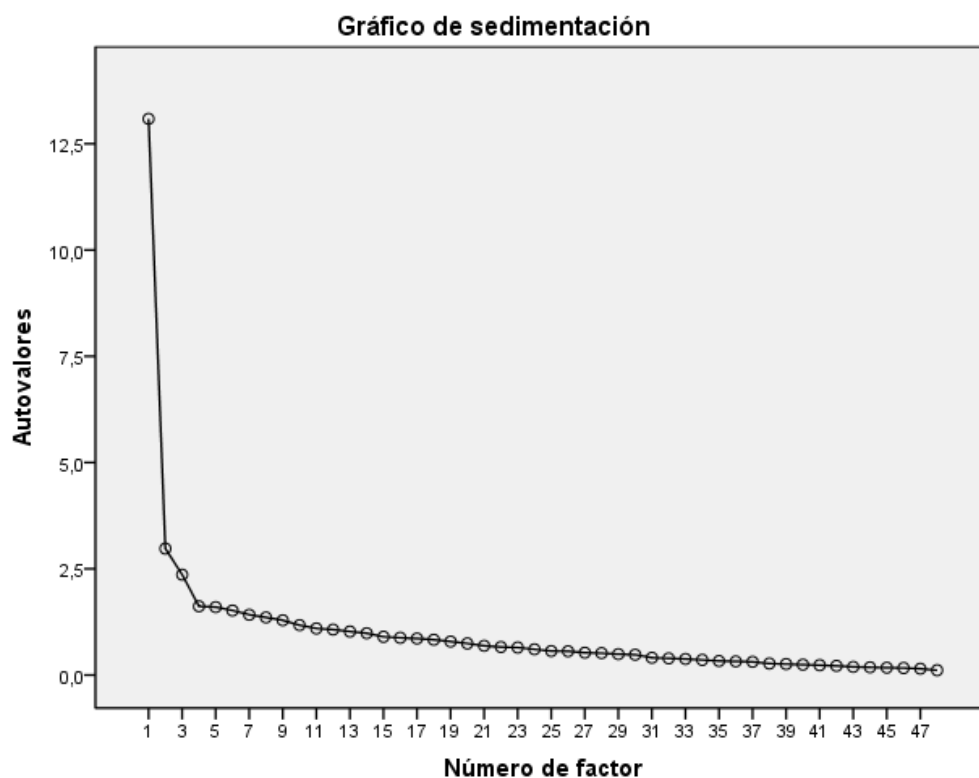


Tabla 7.5. Análisis factorial conjunto de los ítems del BDI-FS, SDS y BAI: matriz de configuración de la solución trifactorial

<i>Ítem</i>	<i>Instrumento</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>
1. Tristeza	BDI-FS	-	,540	-
2. Pesimismo	BDI-FS	-	,710	,261
3. Sentimientos de fracaso	BDI-FS	-	,679	-
4. Pérdida de placer	BDI-FS	-	,664	-
5. Insatisfacción con uno mismo	BDI-FS	-	,514	-
6. Autocríticas	BDI-FS	-	,356	-
7. Pensamientos de suicidio	BDI-FS	-	-	-
1. Tristeza	SDS	-	,617	-
2. Mejoría matutina	SDS	-	-	-
3. Llanto	SDS	-	,545	-
4. Problemas de sueño	SDS	-	,365	-
5. Sin cambios de apetito	SDS	-	-	,432
6. Sin cambios en interés por sexo	SDS	-	,285	-
7. Pérdida de peso	SDS	-	-	-
8. Estreñimiento	SDS	-	,436	-
9. Palpitaciones	SDS	,439	,301	-
10. Cansancio	SDS		,765	-
11. Cabeza despejada	SDS	,271	-	,365
12. Hacer las cosas con misma facilidad	SDS	-	,300	,327
13. Agitación e intranquilidad	SDS	-	-	-
14. Esperanza y confianza en el futuro	SDS	-	-	,608
15. Irritabilidad	SDS	,251	-	-
16. Facilidad para tomar decisiones	SDS	-	-	,588
17. Sentimientos de ser útil	SDS	-	-	,696
18. Sentimientos de vida plena-agradable	SDS	-	-	,443
19. Ideación suicida	SDS	-	,336	-

Nota. *N* = 181. No se muestran las saturaciones < 0,25 y las saturaciones > 0,40 se muestran en negrita.

Tabla 7.5. Análisis factorial conjunto de los ítems del BDI-FS, SDS y BAI: matriz de configuración de la solución trifactorial (Continuación)

<i>Ítem</i>	<i>Instrumento</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>
20. Placer y disfrute con mismas cosas	SDS	-	-	,631
1. Hormigueo o entumecimiento	BAI	,442	-	-
2. Sensación de calor	BAI	,472	-	-
3. Debilidad en las piernas	BAI	-	,562	-
4. Incapacidad para relajarme	BAI	,566	-	-
5. Miedo a que suceda lo peor	BAI	,515	-	-
6. Mareos o vértigos	BAI	,458	-	-
7. Palpitaciones o taquicardia	BAI	,415	,285	-
8. Sensación de inestabilidad	BAI	,607	-	-
9. Sensación de estar aterrorizado	BAI	,307	-	-
10. Nerviosismo	BAI	,403	-	-
11. Sensación de ahogo	BAI	,636	-	-
12. Temblor de manos	BAI	,630	-	-
13. Temblor generalizado	BAI	,831	-	-
14. Miedo a perder el control	BAI	,554	-	-
15. Dificultad para respirar	BAI	,631	-	-
16. Miedo a morir	BAI	,493	-	-
17. Estar asustado	BAI	,719	-	-
18. Indigestión o molestia abdominal	BAI	,428	-	-
19. Sensación de desmayarse	BAI	,810	-	-
20. Rubor facial	BAI	,699	-	-
21. Sudoración	BAI	,568	-	-

Nota. *N* = 181. No se muestran las saturaciones < 0,25 y las saturaciones > 0,40 se muestran en negrita.

En la Tabla 7.5 se presenta la matriz de configuración correspondiente a la solución trifactorial de los ítems del BDI-FS, SDS y BAI. Como puede observarse en dicha tabla, la matriz de configuración confirmó de nuevo la existencia de factores diferentes para los constructos de depresión y ansiedad, y también confirmó de nuevo que el BDI-FS definía de forma sustantiva un

factor de depresión. Efectivamente, considerando de nuevo como despreciables las saturaciones menores de 0,25 y como definitorias las saturaciones iguales o mayores de 0,40, el primer factor de la solución trifactorial era claramente un factor de ansiedad, ya que todos los ítems que lo definían, salvo uno, pertenecían al BAI, de manera que 19 de sus 21 ítems presentaban saturaciones mayores de 0,40 en ese primer factor. El único ítem que no pertenecía al BAI y que presentaba una saturación mayor de 0,40 en ese primer factor era, de nuevo, el ítem de Palpitaciones de la SDS que, como ya se dijo, es perfectamente compatible con la clínica ansiosa. Por su parte, parece también claro que los factores segundo y tercero de la solución trifactorial eran dos factores de depresión, ya que todos los ítems que los definían pertenecían, salvo uno, al BDI-FS o a la SDS. De hecho, 5 de los 7 ítems del BDI-FS y 4 de los 20 ítems del SDS presentaban saturaciones mayores de 0,40 en el segundo factor, y 6 de los restantes 14 ítems del SDS presentaban saturaciones mayores de 0,40 en el tercer factor.

Cabría preguntarse, no obstante, el motivo de la existencia de dos factores de depresión en la muestra de pacientes de AP del presente estudio ya que la misma podría poner en cuestión la validez factorial del BDI-FS, máxime cuando parece que ambos factores no se distinguen mucho por el contenido de los síntomas depresivos que incluyen, ya que, por ejemplo, en ambos factores hay un importante número de síntomas cognitivos definitorios. En concreto, los ítems del BDI-FS de Pesimismo, Sentimiento de fracaso e Insatisfacción con uno mismo muestran saturaciones importantes en el Factor 2, pero, con un contenido sintomatológico muy relacionado, los ítems del SDS de Esperanza y confianza en el futuro, Sentimientos de vida plena y Sentimiento de ser útil muestran también saturaciones muy altas, pero en el

Factor 3. Sin embargo, la existencia de esos dos factores de depresión parece cuestionar más la validez factorial de la SDS que la del BDI-FS, ya que el tercer factor parece reflejar la existencia de un “factor de método” en la SDS.

Un “factor de método” sería aquél que viene determinado más por el método de medida que por el contenido de los ítems, es decir, más por algún factor asociado al formato de los ítems que por el constructo que pretende medir el ítem. En este sentido, varios estudios que han analizado factorialmente instrumentos que incluyen ítems directos (o formulados positivamente) e ítems inversos (o formulados negativamente) han encontrado factores de método asociados a los ítems inversos, es decir, han informado de la aparición de dos factores, uno definido por los ítems directos y otro definido por los ítems inversos, cuando tales instrumentos no pretenden medir constructos múltiples y, además, no hay otros datos que indiquen que esos constructos son múltiples (p. ej., Marsh, 1996; Spector, Van Katwyk, Brannick y Chen, 1997; Tomás, Meléndez, Oliver, Navarro y Zaragoza, 2010; Van der Heiden, Muris, Bos y van der Molen, 2010). Coherentemente con los resultados de esos estudios, todos los ítems que saturaron por encima de 0,40 en el tercer factor fueron ítems inversos de la SDS e, incluso, los ítems que saturaron por encima de 0,25, pero por debajo de 0,40, en ese tercer factor también fueron ítems inversos de la SDS, de manera que 8 de los 10 ítems inversos de la SDS tuvieron saturaciones definitorias ($\geq 0,40$) o no despreciables (0,25-0,39) en el Factor 3.

Por tanto, aunque no es posible determinar los motivos que explican la presencia de un factor de método asociado a los ítems inversos de la SDS (p. ej., dificultades lectoras para entender los ítems inversos, deseabilidad social,

evaluación de otro constructo más amplio o distinto al de depresión) ni era ése uno de los objetivos del presente estudio, lo que sí parece claro observando el patrón de saturaciones factoriales de la Tabla 7.5 es que el BDI-FS definía un factor de depresión diferenciado del factor de ansiedad definido por el BAI.

Para corroborar de forma aún más clara esta última afirmación, se realizó un tercer análisis factorial únicamente con los ítems del BDI-FS y del BAI, y extrayéndose dos factores mediante ejes principales y rotándolos mediante *promax*. La matriz de configuración de esta solución bifactorial se recoge en la Tabla 7.6, solución que, además, era la recomendada por el gráfico de sedimentación (*scree test*) de Cattell y por el test de mínima correlación parcial promediada (MAP) de Velicer.

Tabla 7.6. Análisis factorial conjunto de los ítems del BDI-FS y BAI: matriz de configuración de la solución bifactorial

Ítem	Instrumento	Factor 1	Factor 2
1. Tristeza	BDI-FS	–	0,518
2. Pesimismo	BDI-FS	-0,262	0,856
3. Sentimientos de fracaso	BDI-FS	–	0,766
4. Pérdida de placer	BDI-FS	–	0,709
5. Insatisfacción con uno mismo	BDI-FS	–	0,574
6. Auto-críticas	BDI-FS	–	0,501
7. Pensamientos o deseos de suicidio	BDI-FS	0,266	0,256
1. Hormigueo o entumecimiento	BAI	0,474	–
2. Sensación de calor	BAI	0,475	–
3. Debilidad en las piernas	BAI	0,288	0,363
4. Incapacidad para relajarme	BAI	0,565	–
5. Miedo a que suceda lo peor	BAI	0,566	–
6. Mareos o vértigos	BAI	0,434	–
7. Palpitaciones o taquicardia	BAI	0,427	–
8. Sensación de inestabilidad	BAI	0,620	–
9. Sensación de estar aterrorizado	BAI	0,307	–
10. Nerviosismo	BAI	0,406	0,308
11. Sensación de ahogo	BAI	0,638	–
12. Temblor de manos	BAI	0,603	–
13. Temblor generalizado	BAI	0,816	–
14. Miedo a perder el control	BAI	0,581	–
15. Dificultad para respirar	BAI	0,625	–
16. Miedo a morir	BAI	0,534	–
17. Estar asustado	BAI	0,738	–
18. Indigestión o molestia abdominal	BAI	0,435	–
19. Sensación de desmayarse	BAI	0,823	-0,309
20. Rubor facial	BAI	0,624	–
21. Sudoración	BAI	0,563	–

Nota. N = 190. No se muestran las saturaciones < 0,25 y las saturaciones > 0,40 se muestran en negrita.

Como puede observarse en la Tabla 7.6, la matriz de configuración resultante de la solución bifactorial del BDI-FS y del BAI reflejó claramente la existencia de un factor de ansiedad y otro de depresión, los cuales correlacionaban entre sí de forma alta ($r = 0,60$), pero quedaban claramente definidos por los ítems de sus correspondientes instrumentos.

7.3.5. Análisis factorial

Para conocer con mayor profundidad la estructura factorial del BDI-FS en esta muestra de pacientes de AP se realizó un análisis factorial sobre las respuestas a sus ítems. Previamente se había analizado la viabilidad de dicho análisis a través del índice de Kaiser-Meyer-Olkin (KMO) y de la prueba de esfericidad de Bartlett. El KMO arrojó un valor de 0,88, valor que superaba el estándar de 0,60 que se considera el mínimo para estimar una buena adecuación de los datos muestrales al análisis factorial. La prueba de esfericidad de Bartlett fue estadísticamente significativa [$\chi^2 (21, N = 198) = 1057.038, p < 0,001$], con lo que se rechazó la hipótesis nula de que la matriz de correlaciones entre los ítems del BDI-FS era una matriz identidad, es decir, la hipótesis nula de la no existencia de correlaciones entre los ítems del BDI-FS.

El análisis factorial realizado mediante el procedimiento de factorización de ejes principales reveló la existencia de una solución unifactorial para el BDI-FS, ya que dicho análisis extrajo un solo factor con un autovalor superior a 1 (criterio de Kaiser; véase la Tabla 7.7), el gráfico de sedimentación (*scree test*) de Cattell indicaba la existencia de un solo factor (véase la Figura 7.4), y el test

de mínima correlación parcial promediada (MAP) de Velicer también recomendaba la extracción de un único factor.

Tabla 7.7. Autovalores y porcentajes de varianza de los factores extraídos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 2º

<i>Factores</i>	<i>Autovalor</i>	<i>Porcentaje de la varianza</i>
1	3,285	46,93
2	,946	13,52
3	,823	11,75
4	,573	8,18
5	,514	7,34
6	,477	6,81
7	,382	5,46

Nota. *N* = 198

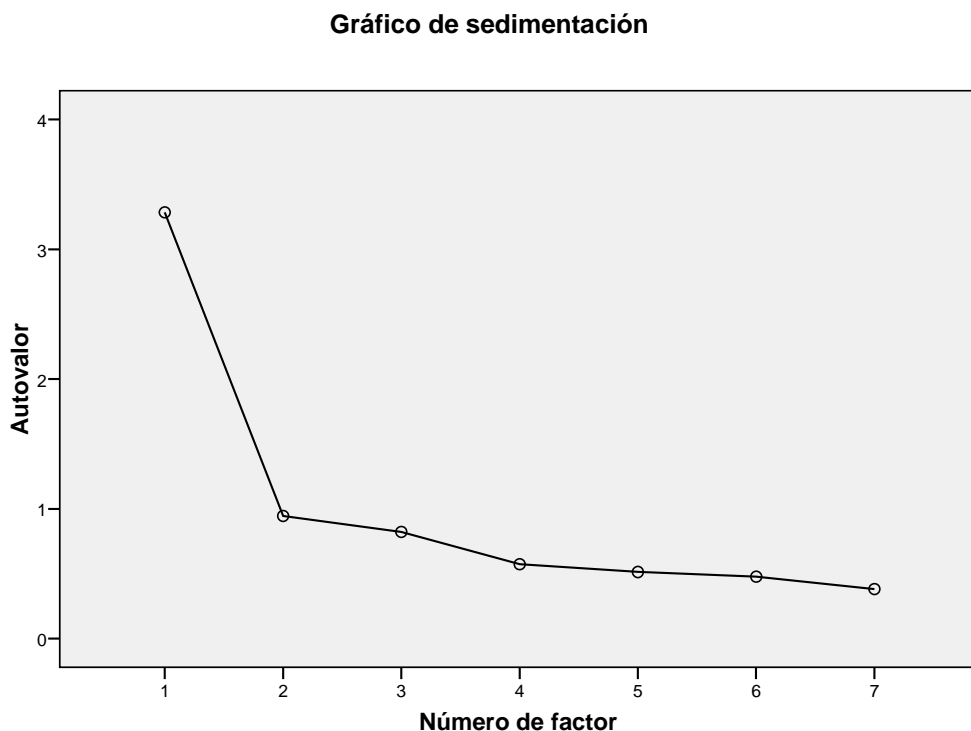
Tabla 7.8. Matriz factorial del BDI-SF en la muestra de pacientes del Estudio 2º tras la extracción de un factor mediante el método de ejes principales

<i>Ítem del BDI-FS</i>	<i>Factor 1</i>
Tristeza	,635
Pérdida de placer	,698
Insatisfacción con uno mismo	,707
Sentimientos de fracaso	,668
Pesimismo	,595
Pensamientos o deseos de suicidio	,566
Autocríticas	,428

Nota. *N* = 198

Como puede verse en la Tabla 7.7, el factor extraído tenía un autovalor igual a 3,28 que explicaba casi el 47% de la varianza de las respuestas de los pacientes a los ítems del BDI-FS, mientras que el resto de factores explicaban individualmente porcentajes de varianza que no llegaban al 14% de varianza. De hecho, la matriz factorial de la solución unifactorial también sugería su plausibilidad, ya que todos los ítems del BDI-FS saturaban 0,40 o más en ese único factor (véase la Tabla 7.8).

Figura 7.4. Gráfico de sedimentación de Cattell de los factores obtenidos mediante factorización de ejes principales de las respuestas a los ítems del BDI-FS de la muestra de pacientes de atención primaria del Estudio 2º



7.3.6. Diferencias en el BDI-FS en función de las características demográficas

En la Tabla 7.9 se presentan, en función de las características demográficas más importantes (sexo, edad, estado civil y nivel de estudios), las medias y desviaciones típicas de las puntuaciones en el BDI-FS de los pacientes de AP del presente estudio.

El ANCOVA realizado sobre dichas puntuaciones tomando como factor intersujetos el sexo y como covariables el resto de variables demográficas reveló que, independientemente de la edad, el estado civil y el nivel de estudios, las mujeres puntuaban significativamente más alto en el BDI-FS que los varones [media ajustada = 3,5 frente a 2,3; $F(1, 453) = 14,31, p = 0,001$].

Un ANCOVA parecido tomando como factor intersujetos la edad y como covariables el sexo, el estado civil y el nivel de estudios también resultó en un efecto estadísticamente significativo del factor edad sobre las puntuaciones del BDI-FS [$F(3, 451) = 4,87, p = 0,002$] que era independiente de las diferencias individuales en las restantes variables demográficas. Las posteriores pruebas *t* de Student sobre las medias ajustadas (véase la Tabla 7.9) y con los niveles de significación corregidos según el procedimiento de Bonferroni, indicaron que ese efecto de la edad quedaba reflejado en una diferencia estadísticamente significativa entre los pacientes de 16-29 años y los pacientes de 45-59 años, de manera que estos últimos puntuaban más alto en el BDI-FS que los primeros (media ajustada = 3,9 frente a 2,04, significativa con $p < 0,05$), sin que ninguna otra diferencia de medias entre grupos de edad en el BDI-FS fuera estadísticamente significativa.

Por el contrario, el ANCOVA con el nivel de estudios como factor intersujetos y el sexo, la edad y el estado civil como covariables, no reveló ningún efecto estadísticamente significativo del nivel de estudios sobre las puntuaciones en el BDI-FS [$F(3, 451) = 2,06, p = 0,105$] (véase la Tabla 7.9).

Tabla 7.9. Puntuaciones medias, desviaciones típicas (DT) y puntuaciones medias ajustadas en el BDI-FS de la muestra conjunta de los Estudios 1º y 2º en función de sus características demográficas y resultados de los ANCOVA unifactoriales sobre las puntuaciones en el BDI-FS

Característica demográfica	N	Media	DT	Medias ajustadas*	ANCOVA [†]	
					p	η ²
Sexo:					0,001	0,016
- Mujeres	275	3,50	3,73	3,47		
- Varones	186	2,24	2,63	2,29		
Edad:					0,002	0,016
- 18-29 años	112	2,32	2,62	2,04		
- 30-44 años	172	2,67	3,06	2,83		
- 45-59 años	105	3,72	3,77	3,90		
- 60 años o más	72	3,75	4,22	3,54		
Estado civil:					0,009	0,011
- Soltero	130	2,76	3,24	3,28		
- Casado o conviviendo con pareja de forma estable	284	2,80	3,08	2,66		
- Separado, divorciado o viudo	47	4,83	4, 83	4,18		
Nivel de estudios:					0,105	0,007
- No tiene estudios	20	3,10	3,94	1,59		
- Estudios primarios, EGB o equivalente	166	3,57	3,67	3,36		
- Bachillerato, FP o equivalente	170	2,71	3,12	2,95		
- Estudios universitarios	103	2,48	3,15	2,74		

Nota. *Medias ajustadas en función de las otras características demográficas (covariables) de la tabla. [†]ANCOVA unifactoriales considerando, sucesivamente, cada característica demográfica como factor intersujetos y el restos de características demográficas como covariables.

Finalmente, en relación con el estado civil, el correspondiente ANCOVA sobre las puntuaciones en el BDI-FS tomando como factor intersujetos dicha variable demográfica y como covariables las restantes reveló que, independientemente del sexo, la edad y el nivel de estudios, existían diferencias estadísticamente significativas en la puntuación media en el BDI-FS en función del estado civil de los pacientes de AP [$F(2, 453) = 4,79, p = 0,009$]. En concreto, las posteriores pruebas t sobre las medias ajustadas revelaron que los pacientes separados, divorciados o viudos puntuaban más alto en el BDI-FS que los pacientes casados o que convivían con una pareja de manera estable (media ajustada = 4,2 frente a 2,7, significativa con $p < 0,05$), sin que ninguna otra diferencia de medias en el BDI-FS entre distintas categorías de estado civil fuera estadísticamente significativa (véase la Tabla 7.9).

En resumen, el sexo, la edad y el estado civil, pero no el nivel de estudios, mostraron efectos significativos independientes sobre las puntuaciones en el BDI-FS, aunque, tal y como se muestra en la Tabla 7.9, los tamaños de dichos efectos fueron pequeños ($\eta^2 \leq 0,016$), situándose lejos de los valores que según los estándares de Cohen (1988) se consideran un tamaño del efecto medio ($\eta^2 \geq 0,0588$).

7.3.7. Baremos

En la Tabla 7.10 se recogen, tanto para la muestra total conjunta de pacientes de AP como para dicha muestra subdivida en mujeres y varones, los percentiles más habituales, la media y la desviación típica de las puntuaciones del BDI-FS. Estos estadísticos pueden servir como baremos para interpretar la puntuación en el BDI-FS de un paciente en concreto.

Dadas las diferencias significativas en el BDI-FS que los análisis anteriores revelaron respecto al sexo, la edad y el estado civil, cabría plantearse la necesidad de calcular esos baremos de manera diferenciada no sólo para el sexo, sino también para cada una de las categorías de las otras variables sociodemográficas. Sin embargo, de acuerdo a los criterios propuestos por Prieto y Muñiz (2000) para los tamaños de los grupos de participantes necesarios para elaborar unos baremos, los grupos de pacientes correspondientes a esas categorías demográficas no llegaron en algunos casos a ser suficientes según esos criterios ($150 \leq N < 300$). La excepción en este sentido fue el sexo, ya que tanto el grupo de pacientes varones como el grupo de pacientes mujeres superaban el número mínimo de 150 personas y, por esa razón, los percentiles, medias y desviaciones típicas de la Tabla 7.10 también se han calculado de forma separada en varones y mujeres.

Tabla 7.10. Baremos del BDI-FS para pacientes españoles de atención primaria obtenidos a partir de la muestra conjunta de los Estudios 1º y 2º

<i>Percentil</i>	<i>Muestra total</i>	<i>Varones</i>	<i>Mujeres</i>
1	0	0	0
5	0	0	0
10	0	0	0
20	0	0	4
25	0	0	1
30	1	0	1
40	1	1	1
50	2	2	2
60	3	2	3
70	4	3	4
75	4	3	5
80	5	4	6
90	7	5	9
95	10	7	12
99	16	12	16
Media	2,92	2,18	3,43
DT	3,34	2,57	3,71
N	493	198	291

7.4. Discusión

El objetivo principal de este Estudio 2º fue profundizar en el análisis de las propiedades psicométricas del BDI-FS en muestras españolas de pacientes de AP y, en general, sus resultados fueron muy favorables en cuanto a la calidad psicométrica del BDI-FS, presentando índices muy satisfactorios respecto a características psicométricas tan importantes como la fiabilidad de consistencia interna, la fiabilidad test-retest, la validez convergente, la validez discriminante y la validez factorial. En este sentido, los resultados de este estudio con una muestra de pacientes de AP procedentes de otra ciudad española distinta a la ciudad en la que se realizó el Estudio 1º replican las buenas propiedades psicométricas del BDI-FS encontradas en ese primer estudio y además aportan nuevos datos de fiabilidad test-retest y de validez convergente y discriminante que también resultaron ser positivos.

7.4.1 Sensibilidad del BDI-FS

Al igual que ocurrió en el Estudio 1º y al igual que han informado estudios previos con muestras de pacientes no psicopatológicos (p. ej., Beck et al., 2000), las puntuaciones del BDI-FS en la muestra de pacientes de AP del Estudio 2º no se distribuyeron de forma normal, sino en forma de una curva leptocúrtica desplazada a la izquierda, lo cual, como se discutió en el capítulo anterior, es coherente con la naturaleza del BDI-FS y con sus características cuando se utiliza en muestras de pacientes de AP.

De hecho, los indicadores de tendencia central y variabilidad de la distribución de las puntuaciones del BDI-FS en la presente muestra de pacientes de AP se asemejaron a los encontrados en el Estudio 1º y en estudios previos en muestras de pacientes médicos de otros países. Así, en la

muestra de pacientes de AP de este estudio, la puntuación media del BDI-FS y su desviación típica fueron, respectivamente, 3,02 y 3,16, datos que son compatibles con los encontrados en el Estudio 1º (media = 2,85 y $DT = 3,47$) y a su vez consonantes con los encontrado en otros estudios con pacientes ambulatorios no psiquiátricos (p. ej., media = 2,18 y $DT = 2,96$ en Steer y Thomas, 1999; media = 2,33 y $DT = 3,13$ en Scheinthal et al., 2001).

Si se analiza la distribución de puntuaciones del BDI-FS por ítems, en ambos estudios (1º y 2º) coinciden las puntuaciones medias máximas en el ítem de Autocrítica (0,73 y 0,72 respectivamente) y también las puntuaciones medias mínimas en el ítem de Suicidio (0,11 y 0,12 respectivamente). Es curioso, en este sentido, que la autocrítica, que no es un síntoma puramente patognomónico de la depresión, aparezca con tanta frecuencia e intensidad entre los pacientes con AP, lo cual podría hacer pensar que dicho síntoma apela a un malestar generalizado y más fácil de expresar entre dichos pacientes.

Respecto al ítem de Suicidio (ítem 7), ya se ha comentado en el capítulo anterior que existe una cierta controversia en la literatura científica sobre si sería conveniente su eliminación del BDI-FS debido a que en algún estudio ha presentado pobres índices de consistencia interna (Golden et al., 2007) o a que algunos investigadores y clínicos consideran que es una pregunta demasiado “directa o fuerte” para personas que están enfermas (Wilheim et al., 2004).

Es cierto que, teniendo en cuenta conjuntamente los datos de los Estudios 1º y 2º, el ítem de Suicidio apenas ha sido puntuado por un 10,5% de los pacientes de AP y prácticamente en todos esos casos (90,4%) en su menor grado de puntuación (1 = “Tengo pensamientos de suicidio, pero no los llevaría

a cabo”), pero, si se analizan los datos de la submuestra de pacientes de AP a los que en el Estudio 1º se les administró la SCID-I, más de la mitad (50,5%) de los pacientes de AP que sufrían un trastorno depresivo mayor puntuaron en dicho ítem, mientras que entre los pacientes de AP que no sufrían tal trastorno, ese porcentaje fue de tan sólo el 2,5%, por lo que dicho ítem puede ser muy discriminativo a la hora de valorar la posibilidad de un trastorno depresivo, máxime si se puntúa en el mismo 2 ó 3. Además, de acuerdo con Beck et al. (2000, p. 4), el ítem de Suicidio debe valorarse en sí mismo, más allá de la puntuación global del instrumento, ya que en conjunción con el ítem de Pesimismo (ítem 2) es un indicador de posible riesgo de suicidio. Por otra parte, tanto en este segundo estudio como en el primero, el ítem de Suicidio mostró entre los pacientes españoles de AP índices de consistencia interna (correlación corregida ítem-total) adecuados que superaban los estándares de 0,30 (Nunnally y Bernstein, 1995) o 0,35 (Cohen y Manion, 1990) que se han propuesto en la literatura científica. En conclusión, pues, los datos de este estudio y del Estudio 1º no apoyan la conveniencia de eliminar el ítem de Suicidio del BDI-FS, sino que, por el contrario, sugieren que el mismo es un elemento sustancial del instrumento que puede ser muy útil en la práctica clínica para detectar casos de depresión mayor, para evaluar el riesgo de suicidio y, en general, para evaluar la sintomatología depresiva de los pacientes de AP.

7.4.2. Fiabilidad del BDI-FS

En el presente estudio se examinó la fiabilidad del BDI-FS en pacientes de AP tanto desde la perspectiva de la consistencia interna como desde la

perspectiva de la estabilidad temporal. En relación con la fiabilidad de consistencia interna, en el presente estudio se obtuvieron índices, tanto en términos del coeficiente *alfa* de Cronbach (igual a 0,80) como en términos de la correlación media entre ítems (igual a 0,37), ligeramente inferiores a los encontrados en el Estudio 1º (0,85 y 0,45, respectivamente). No obstante, teniendo en cuenta los estándares actuales para valorar la fiabilidad de consistencia interna (Bollen y Lennox, 1991; Briggs y Cheek, 1986; Prieto y Muñoz, 2000), los resultados encontrados en el presente estudio también indican, como lo hacían los resultados del Estudio 1º, que la versión española del BDI-FS tiene buenos índices de fiabilidad de consistencia interna en pacientes de AP.

Por ejemplo, Prieto y Muñoz (2000) han propuesto que coeficientes de consistencia interna iguales o mayores de 0,80, pero menores que 0,85, pueden considerarse buenos, mientras que coeficientes iguales o mayores de 0,85 se pueden considerar excelentes. Los coeficientes *alfa* de Cronbach encontrados en este estudio y en el Estudio 1º (0,80 y 0,85), permiten situar a la versión española del BDI-FS en su aplicación en pacientes de AP en niveles de consistencia interna buenos-excelentes, lo cual es consistente con los resultados obtenidos en estudios previos con diversas muestras de pacientes médicos y que informaban, en general, de coeficientes *alfa* en el rango de 0,83 a 0,88 (Beck et al., 1999; Golden et al., 2007; Kefetz et al., 2004; Orive et al., 2010; Steer et al., 1999; Winter et al., 1999; para más información sobre estos estudios, véase la Discusión del Capítulo 6).

En relación con la fiabilidad de estabilidad temporal o fiabilidad test-retest, es importante señalar que este estudio es, junto al estudio Healey et al.

(2008), los dos únicos estudios que han analizado ese tipo de fiabilidad en el BDI-FS y, en concreto, este estudio es el primero que lo ha hecho en una muestra de pacientes de AP, ya que el estudio de Healey et al. (2008) se realizó con una muestra de supervivientes de infarto de corazón. Los resultados al respecto son, en general, prometedores. La correlación test-retest a los 7 días en el presente estudio fue igual a 0,89, mientras que en el estudio de Healey et al. (2008) la correlación test-retest a los 7-10 días fue igual a 0,63. De acuerdo a los criterios propuestos por Prieto y Muñiz (2000) para valorar los coeficientes de estabilidad, esos resultados indican que en muestras de pacientes de AP sería esperable obtener índices excelentes de fiabilidad test-retest a la semana (correlaciones iguales o mayores que 0,80; Prieto y Muñiz, 2000), mientras que en muestras de pacientes con infarto de miocardio sería esperable obtener índices adecuados, pero con algunas carencias (correlaciones mayores o iguales a 0,55, pero menores de 0,65; Prieto y Muñiz, 2000).

En conclusión, pues, en pacientes de AP la estabilidad temporal a la semana de las puntuaciones que ofrece la versión española del BDI-FS es elevada y, de hecho, aunque se encontró en el retest el habitual descenso en las puntuaciones medias que suele caracterizar a los instrumentos psicopatológicos cuando se administran por segunda vez (Bumberry, Oliver y McClure, 1978; Hatzenbuehler, Parpal y Matthews, 1983; Oliver y Burkham, 1979; Piacentini et al., 1999), dicho descenso, aunque estadísticamente significativo, fue muy pequeño en términos del tamaño del efecto (d de Cohen = 0,07) y clínicamente irrelevante (como media no llegó a alcanzar ni un cuarto de punto, es decir, 0,22). De hecho, en el 65,7% de los pacientes la puntuación

del BDI-FS en el retest fue exactamente la misma que una semana antes, y en otro 20,5% de los pacientes la puntuación varió en un solo punto (± 1).

7.4.3 Validez convergente

En cuanto a la validez convergente de la versión española del BDI-FS, en el presente estudio se encontró que el BDI-FS correlacionaba de forma positiva y significativa con la SDS de Zung, un cuestionario que se encuentra entre los instrumentos más utilizados y validados para la evaluación de la sintomatología depresiva en todo tipo de poblaciones, incluyendo la población de pacientes de AP (Feightner y Worrall, 1990; Mulrow et al., 1995). Además, el coeficiente de correlación encontrado ($r = 0,71$) fue similar a los coeficientes de correlación hallados en el Estudio 1º tanto en respecto a otro cuestionario de depresión, la subescala de depresión de la HAD, como respecto a una medida heteroevaluada de depresión, el diagnóstico de trastorno depresivo mayor realizado por un clínico a partir de la entrevista estructurada SCID-I ($r = 0,73$ y $0,76$, respectivamente), y todos ellos, según los criterios de Prieto y Muñiz (2000), pueden considerarse índices de validez convergente excelentes ($r \geq 0,60$).

Es más, el coeficiente de validez convergente para el BDI-FS obtenido en el presente estudio es consonante con los coeficientes hallados en otros estudios nacionales e internacionales y que también respaldan la buena validez convergente del BDI-FS (véase la Discusión del Capítulo 6), por lo que se podría concluir que la versión española del BDI-FS es un instrumento que mide fundamentalmente sintomatología depresiva o, al menos, mide el mismo

constructo que los instrumentos que actualmente se consideran medidas de referencia de la depresión (p. ej., SDS, SCID-I o HAD).

7.4.4 Validez discriminante

Para evaluar la validez discriminante del BDI-FS, en el presente estudio se eligió a la ansiedad como constructo de contraste, al igual que se hizo en el Estudio 1º, aunque en esta ocasión frente a otra medida de la ansiedad, el BAI, uno de los instrumentos de evaluación de la ansiedad más utilizado en la investigación psicológica (Piotrowski y Gallant, 2009). Sin duda, la elección de la ansiedad como constructo de contraste supone poner a prueba la validez discriminante del BDI-FS en las condiciones más exigentes, ya que dado el solapamiento conceptual y empírico entre los constructos de ansiedad y depresión (Sanz, 1991, 1992, 1993; Sanz y Navarro, 2003) cabría esperar que el BDI-FS y el BAI presentaran relaciones significativas que confundieran en alguna medida su validez discriminante. En este sentido, hubiera sido más fácil examinar la validez discriminante del BDI-FS frente a medidas de otros constructos con los que la depresión guarda menos relación como, por ejemplo, ciertos rasgos de personalidad como apertura a la experiencia o amabilidad. No obstante, en la medida que la distinción entre ansiedad y depresión sigue siendo válida y útil en psicopatología tanto a nivel sintomático o sindrómico como a nivel nosológico (sin embargo, véanse los capítulos de Goldberg, Kendler, Sirovatka y Regier, 2011, para una visión crítica de esa distinción, especialmente entre trastorno de ansiedad generalizada y trastorno depresivo mayor), no cabe duda de que examinar la validez discriminante de los instrumentos de depresión respecto a las medidas de ansiedad, o

viceversa, es un objetivo importante a la hora de evaluar la calidad psicométrica de dichos instrumentos.

En resumen, examinar la capacidad de los instrumentos de depresión (o ansiedad) para discriminar entre depresión y ansiedad es una tarea ardua y controvertida, puesto que recala en la complejidad de los constructos latentes, es decir, en la propia naturaleza del constructo depresión, difícilmente dissociable del de ansiedad. En este sentido, en el presente estudio, tal y como cabría esperar, se encontró una correlación positiva y significativa entre el BDI-FS y el BAI igual a 0,57, correlación que fue algo menor a la encontrada entre el BDI-FS y la subescala de ansiedad de la HAD en el Estudio 1º ($r = 0,56$) o entre el BDI-II y el BAI en población general española ($r = 0,63$; Magán et al., 2008), pero similar a la encontrada entre esos dos últimos instrumentos en muestras españolas de estudiantes universitarios ($r = 0,58$; Sanz y Navarro, 2003) y de pacientes psicopatológicos ($r = 0,56$; Sanz et al., 2011).

Aunque la magnitud grande del coeficiente de correlación entre el BDI-FS y el BAI puede lógicamente suscitar alguna duda sobre la validez discriminante del BDI-FS, en el presente estudio se encontraron otros dos indicios importantes que avalan dicha validez discriminante en la misma dirección que se encontró en el Estudio 1º, aunque respecto a otros instrumentos de ansiedad y depresión. En primer lugar, la correlación que mostró el BDI-FS con el BAI en este estudio fue significativamente menor que la correlación que mostró con la SDS ($r = 0,57$ frente a 0,71). En segundo lugar, diversos análisis factoriales con los ítems conjuntos del BDI-FS, el BAI y la SDS, y con soluciones bifactoriales y trifactoriales, revelaron de manera consistente la existencia de factores diferentes, aunque correlacionados, para

los constructos de depresión y ansiedad, y que prácticamente todos los ítems del BDI-FS definían de forma sustantiva el factor de depresión, pero no mostraban saturaciones apreciables en el factor de ansiedad.

Por otro lado, en esta ardua tarea de discriminar depresión y ansiedad, los resultados de validez discriminante obtenidos para el BDI-FS superaban a los alcanzados en este estudio por otra medida similar de depresión, la SDS. En concreto, en comparación con la SDS que correlacionó 0,66 con el BAI, el BDI-FS correlacionó significativamente menos con el BAI ($r = 0,57$ frente 0,66). Igualmente, los análisis factoriales de los ítems conjuntos del BDI-FS, el BAI y la SDS mostraron que los ítems del BDI-FS, en comparación a los ítems de la SDS, mostraban un patrón más claro y consistente de saturaciones sustantivas en el factor de depresión y de saturaciones no sustantivas en el factor de ansiedad. Así, por ejemplo, en la solución bifactorial de los tres instrumentos (un factor de ansiedad y otro de depresión), dicho patrón deseable de saturaciones aparecía en 6 de los 7 ítems del BDI-FS (85,7% de sus ítems), pero tan sólo en 9 de los 20 ítems de la SDS (45% de sus ítems), mientras que en la solución trifactorial de los tres instrumentos (un factor de ansiedad y dos factores de depresión), ese patrón de saturaciones aparecía en 5 de los 7 ítems del BDI-FS (71,4% de sus ítems), pero tan sólo en 10 de los 20 ítems de la SDS (50% de sus ítems). Es más, esta última solución trifactorial también sugería que el segundo factor de depresión, en el que sólo saturaban significativamente ítems de la SDS, más que ser solamente un factor sustantivo de depresión, podría ser también un “factor de método”, ya que estaba definido por los ítems inversos de la SDS. Si futuras investigaciones confirmaran la presencia de ese “factor de método”, esto podría suponer que la medida de depresión de la SDS podría estar distorsionada en parte por variables

asociadas a las respuestas a los ítems inversos (p. ej., dificultades lectoras para entender los ítems inversos, deseabilidad social, evaluación de otro constructo más amplio o distinto al de depresión; véase Marsh, 1996; Spector et al., 1997; Tomás et al., 2010; van der Heiden et al., 2010). En consecuencia, pues, cabría concluir que, en pacientes de AP, el BDI-FS parece discriminar mejor la depresión respecto a la ansiedad medida por el BAI de lo que lo hace la SDS.

7.4.5. Validez factorial

En consonancia con la estructura factorial del BDI-FS original (Beck et al. 2000), los resultados de los análisis factoriales realizados en el presente estudio también han encontrado claramente una solución factorial de un único factor para la versión española del BDI-FS, replicando así la solución unifactorial encontrada en el Estudio 1º, pero con una muestra de pacientes españoles de AP procedente de una ciudad distinta a la ciudad en la que se realizó el Estudio 1º. Por supuesto, tal y como se comentó en el capítulo anterior, ese único factor que parece subyacer al BDI-FS justifica la obtención de una puntuación total en el BDI-FS y, además, refleja una dimensión cognitivo-afectiva de la depresión, ya que los ítems del BDI-FS proceden de los ítems del BDI-II que *a priori* evalúan síntomas cognitivos y afectivos y que, consistentemente, forman una dimensión cognitivo-afectiva en los propios análisis factoriales del BDI-II (Arnau et al., 2001; Beck et al., 1996; Buttonm, Grothe y Jones, 2004; Dozois et al., 1998; Sanz, Navarro y Vázquez, 2003; Sanz, Perdigón y Vázquez, 2003; Sanz et al., 2005; Sanz y García-Vera, 2009, 2011; Steer et al., 1999, 2000; Penley et al., 2003; Wiebe y Penley, 2005).

7.4.6. Diferencias en el BDI-FS en función de las características demográficas.

Analizando los datos combinados de la muestra de pacientes del Estudio 1º y de este estudio, se han encontrado algunas diferencias significativas en las puntuaciones del BDI-FS en función de las características demográficas de los pacientes de AP. En concreto, independientemente de la influencia de la edad, el estado civil y el nivel de estudios, las mujeres puntuaron significativamente más alto en el BDI-FS que los varones. Aunque este hallazgo no se corresponde con la ausencia de relación entre sexo y BDI-FS informada por Beck et al. (1997) en una muestra de pacientes de medicina interna ingresados, por Scheinthal et al. (2001) en una muestra de pacientes geriátricos o por Beck et al. (2000) en diferentes muestras de pacientes médicos, sin embargo, es coherente con la literatura científica que de manera muy consistente encuentra mayores tasas de depresión en las mujeres que en los varones en muestras de pacientes de AP (Alonso et al., 2007; Aragonés et al., 2001; Caballero et al., 2008; Gabarrón et al., 2002; Ferrer y Rodríguez, 1999; King et al., 2008; ODIN, 2001; Olivera et al., 2004, 2005, 2008; Serrano et al., 2010).

Analizando los datos combinados de los dos estudios también se han encontrado, tras controlar los efectos del sexo, el estado civil y el nivel de estudios, diferencias significativas en las puntuaciones del BDI-FS de los pacientes de AP en función de la edad. De nuevo, Beck et al. (1997), Scheinthal et al. (2001) y Beck et al (2000) no encontraron una relación de la edad con el BDI-FS en muestras de pacientes médicos y geriátricos estadounidenses, pero, sin embargo, en este estudio con pacientes españoles

de AP, el grupo de pacientes con edades entre 45-59 años mostró, de manera significativa, mayores puntuaciones en el BDI-FS que el grupo de pacientes de menor edad (16-29 años). Esta diferencia es coherente con otros estudios que señalan que los trastornos depresivos son aproximadamente el doble de prevalentes en personas mayores respecto a los más jóvenes (Alexopoulos et al., 1999, citado en Scheinthal et al., 2001), o que muestran en España un aumento en la prevalencia vital de la depresión a medida que aumenta la edad desde la cohorte de 18-34 años hasta llegar a la cohorte de 50-64 años (Gabilondo et al., 2010).

Más allá de los efectos del sexo, la edad o el nivel de estudios, el estado civil también apareció significativamente relacionado con las puntuaciones del BDI-FS en la muestra conjunta de pacientes de AP de los dos estudios de esta tesis doctoral. En concreto, el grupo de pacientes separados, divorciados o viudos obtuvo, de manera significativa, mayores puntuaciones en el BDI-FS que el resto de pacientes, mientras que los pacientes casados o que convivían con una pareja de forma estable obtuvieron, de forma significativa, menores puntuaciones en el BDI-FS que el resto de pacientes. En definitiva, según los resultados del presente estudio, parece que el estar casado o conviviendo con una pareja estable podría ser un factor de protección para la depresión (medida a través de las puntuaciones en el BDI-FS) y, en cambio, el estar divorciado, separado o viudo podría ser un factor de riesgo para la depresión.

Estos resultados son novedosos en cuanto al BDI-FS se refiere, ya que en la literatura científica no se han encontrado hallazgos previos sobre la relación del BDI-FS con el estado civil, pero, en cambio, no lo son en cuanto a la relación entre estado civil y depresión. Por ejemplo, algunos estudios

epidemiológicos realizados en España han encontrado, de forma coincidente a los resultados anteriores, que el haber estado previamente casado (actualmente separado, divorciado o viudo) estaba significativamente relacionado con una mayor prevalencia en el último año de trastornos del estado de ánimo o del episodio depresivo mayor medidos por la entrevista estructurada CIDI (Alonso et al., 2007; Gabilondo et al., 2010; Haro et al., 2005).

Finalmente, respecto al nivel de estudios, no se ha encontrado, tras controlar el efecto de las otras variables demográficas, ninguna relación entre dicha variable y las puntuaciones en el BDI-FS, resultado que replica los hallazgos de Beck et al. (2000) con el BDI-FS en diversas muestras de pacientes médicos, y que también es coincidente con los hallazgos epidemiológicos en España que no encuentran que el nivel de estudios sea una variable significativamente relacionada con una mayor prevalencia en el último año de trastornos del estado de ánimo o del episodio depresivo mayor (Gabilondo et al., 2010; Haro et al., 2006).

7.4.7. Baremos

A partir de los datos combinados de la muestra de pacientes del Estudio 1º y de este estudio, se han elaborado los primeros baremos del BDI-FS para pacientes españoles de AP. Por supuesto, estos baremos tienen las limitaciones inherentes a la naturaleza incidental de las muestras de pacientes con que se han obtenido, pero, al menos en su tamaño, la muestra conjunta de pacientes de AP de los dos estudios ($N = 493$) puede considerarse moderada

($300 \leq N < 600$) para elaborar unos baremos con garantías, según los estándares de Prieto y Muñiz (2000).

Dado que se han encontrado diferencias significativas en el BDI-FS entre grupos de pacientes de distinto sexo, edad o el estado civil, hubiera sido deseable elaborar esos baremos de forma diferenciada para esos grupos. Sin embargo, puesto que los tamaños de esos grupos de pacientes no llegaron en algunos casos a ser suficientes ($N > 150$) para elaborar unos baremos con garantías según los criterios propuestos por Prieto y Muñiz (2000), tan sólo se realizaron dichos baremos de manera separada para hombres y mujeres, para los cuales sí se contaba con grupos de más de 150 pacientes.

Estos baremos deberían utilizarse de manera conjunta con los rangos de puntuaciones correspondientes a distintas categorías de gravedad de la sintomatología depresiva que Beck et al. (2000) han propuesto: depresión mínima o ausente (0-3 puntos), depresión leve (4-6 puntos), depresión moderada (7-9 puntos), y depresión grave (10-21 puntos). De esta manera, se podría obtener una mejor caracterización del nivel de sintomatología depresiva de un paciente concreto. Por ejemplo, aunque puntuaciones de 3 o menos representan una depresión mínima o ausente, los baremos de la Tabla 7.10 sugieren que si un varón puntuase 3 en el BDI-FS, esta puntuación dejaría por debajo al 75% de la muestra normativa correspondiente, mientras que si esa puntuación la obtuviese una mujer, dejaría por debajo tan sólo al 50% de su muestra normativa. Esto querría decir, por ejemplo, que si una mujer que acude a un centro de AP obtiene una puntuación de 3 en el BDI-FS, podemos estar más seguro de que su puntuación está dentro de la normalidad que si esa puntuación la obtuviera un varón.

7.4.8. Limitaciones del Estudio 2º

En resumen, el objetivo principal de este segundo estudio era la replicación de los datos de fiabilidad y validez encontrados en el primer estudio sobre el BDI-FS y la ampliación de dichos datos a propiedades psicométricas no abordadas en el Estudio 1º o a nuevos instrumentos de referencia o contraste no contemplados en las pruebas de validación del Estudio 1º. Los resultados del presente estudio en relación con esos objetivos fueron tan satisfactorios como los encontrados en el primer estudio, de forma que ambos en conjunto respaldan la utilización de la versión española del BDI-FS como un instrumento válido y fiable para la medida de la sintomatología depresiva en pacientes españoles de AP.

Sin embargo, esta conclusión general, como las más específicas que se han venido exponiendo en los epígrafes anteriores, deberían valorarse a la luz de algunas de las limitaciones del presente estudio, limitaciones que son en gran medida parecidas a las que tenía el Estudio 1º y de las cuales, la más importante quizás sea el hecho de que, al igual que ocurría en el primer estudio, la muestra de pacientes de AP que participó en este segundo estudio fue una muestra incidental procedente de los centros de atención primaria de una ciudad española y remitidos por un grupo de conveniencia de médicos. Por tanto, para la selección de la muestra no se siguieron criterios de muestreo aleatorio respecto a otras ciudades españolas o respecto a otros médicos de esos centros de atención primaria. Por todo ello, cabría cuestionarse la validez externa o representatividad de los resultados encontrados. No obstante, sería importante destacar que en el presente estudio, así como en el Estudio 1º, al menos en cuanto a los pacientes seleccionados para cada médico de AP que

accedió a colaborar, se siguió un procedimiento que muestreaba de forma sistemática los pacientes en función de un criterio no sesgado (p. ej., se seleccionaba el paciente correspondiente a cada 15 minutos de la lista de citas de cada médico de AP).

Capítulo 8:

Conclusiones generales

*Es de importancia para quien desee alcanzar
una certeza en su investigación saber dudar a tiempo.*

Aristóteles

La depresión es considerada actualmente, tanto a nivel nacional como internacional, como un problema importante de salud pública. Los principales motivos para esta consideración se han puesto de manifiesto en numerosas ocasiones a lo largo de esta tesis y se podrían resumir en cuatro:

- Una elevada prevalencia en la población, especialmente en la de pacientes de AP. De hecho, la tasa de depresión en personas con enfermedad crónica duplica o triplica a la que se encuentra en personas que gozan de buena salud física (NICE, 2009).
- Un curso crónico. Los últimos estudios longitudinales aportan datos contundentes, la remisión completa y mantenida en el tiempo de la depresión es escasa (Chisholm et al., 2001; Herman et al., 2002), tendiendo algo más de la mitad de las personas con depresión a mantener un curso intermitente o crónico (Stegenga et al., 2010).
- Un grado de interferencia muy alto a nivel personal, familiar y social. Los trastornos depresivos están asociados con un alto grado de disfuncionalidad y pérdida de calidad de vida en la persona que lo padece, sin menoscabo de la elevada relación con múltiples patologías médicas y, en casos extremos, con la mortalidad por suicidio. Los familiares y allegados sufren también en su calidad de vida así como por el aislamiento de la persona con depresión.
- Una importante carga socioeconómica para el sistema sanitario y para la sociedad en general (Karapampa et al., 2011). Se prevé que para el 2030 la depresión unipolar sea la segunda enfermedad en el ranking de mayor carga económica (Curran et al., 2007), pero incluso más

precozmente, en 2020, se estima que la depresión ascenderá a ser la segunda causa de incapacidad e improductividad funcional (Murray et al., 1997). A nivel del sistema sanitario, el gasto farmacéutico y la hiperfrecuentación de consultas tanto en AP como en atención especializada, suponen un importante gasto para las arcas públicas (Binns et al., 2007).

Este dramático panorama actual hace que sean necesarias políticas y protocolos dirigidos a mejorar la detección y monitorización de los cuadros depresivos. Para conseguir estos objetivos aparece como figura fundamental y prioritaria el médico de AP (Lake, 2008; Cepoiu et al., 2008), y como principal hándicap de este colectivo para la consecución de los mismos, las dificultades que presentan para el diagnóstico preciso de los trastornos depresivos, especialmente en relación con el infradiagnóstico, describiéndose de manera sistemática una dificultad de los médicos de AP para reconocer trastornos depresivos presentes en los pacientes, dificultad que se estima entre un tercio y la mitad de los casos con depresión no diagnosticados como tal (Gabarrón et al, 2004; Mitchell et al. 2009; WHO, 2008).

Se han propuesto múltiples causas para estas dificultades, entre otras, la presentación de niveles más leves de gravedad de la depresión entre los pacientes de AP, la no expresión explícita por parte de éstos de los síntomas más puramente depresivos, la confusión sintomática con la patología médica, el escaso tiempo disponible en las consultas, el miedo a abrir la “caja de pandora”, o la baja formación de los médicos de AP en salud mental. Pero, posiblemente, otra de las causas que explicarían estas dificultades es que para la evaluación de los trastornos depresivos, los médicos de AP utilizan casi

exclusivamente la entrevista clínica (Martín-Águeda et al., 2007), en detrimento de otros métodos, como los instrumentos de cribado, que usados complementariamente podrían aumentar el rendimiento diagnóstico de los médicos de AP en cuanto a la detección de la depresión.

Además, en los últimos años ha surgido una creciente preocupación por los falsos positivos en depresión, por lo que puede denominarse la “medicalización de la vida cotidiana”, es decir, por el excesivo diagnóstico de trastornos depresivos, con sus subsecuentes consecuencias negativas, en personas que simplemente sufren reacciones de carácter adaptativo ante un suceso adverso (García-Campayo, 1993; García-Testal et al., 1998, Summerfield y Veale, 2008).

Todas estas dificultades y problemas de diagnóstico en AP se podrían ver mejorados con un instrumento de detección, cribado o despistaje (*screening*) de la depresión que optimizara la información que puede obtener el médico de AP sobre un posible trastorno depresivo. Aunque controvertida, esta recomendación se ve respaldada por las principales guías clínicas para la prevención y tratamiento de la depresión en AP elaboradas por el *US Preventive Services Task Force* (2002), el *Canadian Task Force on Preventive Health Care* (2005) y el *UK National Institute for Health and Clinical Excellence* o NICE (2006), especialmente cuando el uso de un instrumento de cribado de la depresión se implementa dentro de un modelo de cuidados coherente, integrado e individualizado de la persona con trastorno depresivo (Boarman y Walters, 2009).

Como se ha analizado en la presente tesis, actualmente hay disponibles varios instrumentos para el cribado de la depresión con aceptables

propiedades psicométricas, aunque ninguno se utiliza de manera cotidiana en España y sólo unos pocos han sido contruidos de manera específica para AP. El *National Health System* (NHS) del Reino Unido junto con la *British Medical Association* (BMA) de ese país publican periódicamente una guía sobre resultados en investigación y calidad para médicos de AP. En su última guía (2009/2010), recomiendan el uso de tres instrumentos de cribado para la evaluación de la depresión: el *Patient Health Questionnaire* (PHQ-9), la *Hospital Anxiety and Depression Scale* (HAD) y el *Beck Depression Inventory-II* (BDI-II).

Tal y como se ha apuntado anteriormente, una de las dificultades específicas de los médicos de AP para el reconocimiento de la depresión es el solapamiento con síntomas físicos y, precisamente, a partir de uno de los tres instrumentos de cribado recomendados por el NHS y la BMA, el BDI-II, se crea específicamente para AP el Inventario de Depresión de Beck-Cribado Rápido (BDI-FS; Beck et al, 2000), eliminando del BDI-II los síntomas somáticos que pueden dificultar el reconocimiento de la depresión, y, por tanto, obteniéndose un instrumento que facilita la detección de los trastornos depresivos, además de proporcionar una medida de la sintomatología depresiva que permite conocer la gravedad de la depresión de un paciente y monitorizar su posible tratamiento.

Tal y como se ha descrito en capítulos anteriores de esta tesis doctoral, el BDI-FS ha mostrado a nivel internacional unas muy buenas propiedades psicométricas en su uso con pacientes de AP y con otros pacientes médicos. Además cumple todos los requisitos necesarios para ser un buen instrumento

de cribado en AP (OMS, 1968; Muñiz, 2000, 2011; *UK National Screening Committee*, 2011).

A pesar de la importancia de los problemas derivados de la alta prevalencia de la depresión en AP y de su elevado infradiagnóstico, apenas existen estudios científicos en España sobre instrumentos de detección de la depresión específicos para las necesidades de este nivel de asistencia sanitaria. Con la presente tesis doctoral, se pretendía avanzar en el conocimiento de ese ámbito de estudio aportando la única investigación que se sepa sobre la validación del BDI-FS en muestras de pacientes españoles de AP y, en particular, sobre su capacidad para detectar casos de depresión mayor en comparación a la capacidad de los médicos de AP y de otro instrumento de cribado de la depresión utilizado en España.

Esta investigación se ha articulado a partir de dos estudios relacionados. El objetivo principal del primer estudio fue examinar las principales propiedades psicométricas del BDI-FS en una muestra de pacientes españoles de AP y determinar su rendimiento diagnóstico para detectar el trastorno depresivo mayor. Por su parte, el segundo estudio tenía como objetivo principal la replicación de las bondades psicométricas halladas en el primer estudio y la ampliación de la validación del BDI-FS a nuevos índices de fiabilidad y validez no contemplados en el estudio primero, todo lo cual permitiría tener una opinión más sólida sobre sus propiedades psicométricas y, por tanto, sobre su utilidad como instrumento de evaluación y detección de la depresión en AP en España.

A partir de los resultados de ambos estudios se puede afirmar que se han confirmado prácticamente todas las hipótesis planteadas inicialmente sobre los buenos índices de calidad (fiabilidad y validez) del BDI-FS como

instrumento de evaluación y cribado de la depresión, y sobre su mejor rendimiento diagnóstico como instrumento de detección de la depresión frente al juicio clínico del médico de AP o a la subescala de depresión de la HAD. Concretamente, a partir de los resultados de esta investigación se puede concluir que:

1. De manera consistente en dos muestras de pacientes de AP procedentes de dos ciudades de dos comunidades autónomas (Madrid y Castilla y León) distintas, el BDI-FS muestra una solución unifactorial que se corresponde con una dimensión cognitiva-afectiva de la depresión y que **valida factorialmente** la puntuación global del BDI-FS.
2. De manera también consistente en las dos muestras de pacientes de AP, el BDI-FS obtiene índices de **fiabilidad de consistencia interna** $\geq 0,80$, propios de un nivel bueno de fiabilidad de consistencia interna según los estándares actuales.
3. En una muestra de pacientes españoles de AP, el BDI-FS obtiene un índice de **fiabilidad test-retest** a la semana $\geq 0,80$, lo que indica un nivel excelente de fiabilidad de estabilidad temporal según los estándares actuales. Este resultado, amén de consolidar el BDI-FS como un instrumento fiable, es un hallazgo novedoso a nivel internacional, ya que este es el primer estudio que ha analizado esa propiedad psicométrica en pacientes de AP y el único, junto a otro estudio con pacientes cardíacos, que la ha analizado en pacientes médicos de cualquier tipo. De hecho, ni los propios autores del BDI-FS habían analizado esta propiedad psicométrica básica en el manual del instrumento (Beck et al., 2000).

4. Los índices de **validez convergente** del BDI-FS son excelentes en función de los estándares actuales, superando con creces ($\geq 0,70$) el criterio de excelencia ($\geq 0,60$) en dos muestras de pacientes de AP procedentes de dos ciudades distintas y con tres instrumentos de referencia distintos (la SCID-I, la subescala de depresión de la HAD y la SDS), todos ellos empíricamente validados en España para la evaluación de la depresión. Es más, la validez convergente del BDI-FS se ha visto firmemente respaldada por el hecho de que tales instrumentos son tanto de autoinforme (HAD y SDS) como de heteroevaluación (SCID-I), y tanto específicamente creados para la evaluación de la depresión en pacientes médicos (HAD) como para la evaluación de la depresión en pacientes con trastornos mentales (SDS y SCID-I) y, de hecho, uno de ellos (la SCID-I) es un instrumento de referencia para el diagnóstico de los trastornos mentales en general y los trastornos del humor en particular.
5. El BDI-FS correlaciona de forma importante ($\geq 0,50$) y consistente con medidas de ansiedad (subescala de ansiedad de la HAD y BAI) en pacientes españoles de AP, pero tales correlaciones parecen deberse más a las relaciones que guardan los constructos de depresión y ansiedad entre sí que a un problema grave de **validez discriminante** por parte del BDI-FS ya que, por un lado, las correlaciones del BDI-FS con las medidas de ansiedad son en todos los casos menores a las correlaciones convergentes obtenidas con los instrumentos de evaluación de la depresión (subescala de depresión de la HAD, la SCID-I y la SDS) y, por otro lado, diversos análisis factoriales de los ítems

conjuntos del BDI-FS y de los instrumentos de depresión y ansiedad utilizados en cada estudio, revelan de forma consistente en las dos muestras de pacientes españoles de AP soluciones factoriales con factores claramente diferentes de ansiedad y depresión en las cuales los ítems del BDI-FS definen significativa y exclusivamente el factor de depresión. Estos datos sugieren que, a pesar de la relación que guardan los constructos de ansiedad y depresión y a pesar de la correlación que muestra el BDI-FS con instrumentos de ansiedad, el BDI-FS tiene una buena capacidad para discriminar entre los síntomas de depresión y ansiedad.

6. Los índices de **validez de criterio o rendimiento diagnóstico** del BDI-FS para detectar el trastorno depresivo mayor en pacientes de AP son buenos o excelentes (p. ej., área bajo la curva ROC $\geq 0,90$, índices de sensibilidad y especificidad $\geq 0,80$, índice kappa de concordancia con el criterio $\geq 0,80$). Cuando se valida un test, un potente indicador de su calidad es la validez de criterio, pero ésta es fundamental, entendida como rendimiento diagnóstico frente a un criterio diagnóstico de referencia, cuando se trata de validar un instrumento de cribado. A diferencia de gran número de estudios previos que usan como estándar o “patrón oro” otros instrumentos de cribado como, por ejemplo, el PRIME-MD, en la presente investigación se utilizó uno de los criterios de referencia diagnóstico más firmemente consolidados, el diagnóstico realizado por un especialista con experiencia mediante la entrevista estructurada diagnóstica SCID-I.

7. Sea cual sea el índice utilizado (área bajo la curva ROC, sensibilidad, especificidad, índice kappa de concordancia, valores predictivos positivo y negativo, eficiencia, índice de Youden), el rendimiento diagnóstico del BDI-FS para detectar el trastorno depresivo mayor en pacientes de AP es **superior o significativamente superior** al mostrado por la subescala de depresión de la HAD o por el juicio clínico del médico de AP.
8. Aunque el punto de corte habitual del BDI-FS para el cribado de la depresión (puntuación ≥ 4) obtiene buenos índices de rendimiento diagnóstico en pacientes españoles de AP, **aumentar el punto de corte a 6** optimiza en dichos pacientes ese rendimiento obteniendo un 85% de sensibilidad y un 99% de especificidad, y, para distintas estimaciones de la prevalencia del trastorno depresivo (13 o 20%), valores predictivos positivos y negativos por encima de 0,90. No obstante, estos puntos de corte son indicativos, de forma que en la presente tesis doctoral se presentan datos de rendimiento diagnóstico para diferentes puntuaciones del BDI-FS de manera que cada profesional o investigador utilice la que mejor se ajuste a sus necesidades (p. ej., aumentar al máximo la sensibilidad).
9. Existen **diferencias significativas en las puntuaciones del BDI-FS** que obtienen los pacientes españoles de AP en función del sexo (mayor puntuación en las mujeres), la edad (mayor puntuación en los pacientes de 45-59 años frente a los de 16-29 años) y el estado civil (mayor puntuación en los pacientes separados, divorciados o viudos, y menor en los casados o en los que conviven con una pareja estable), pero no

en función del nivel de estudios. Estas diferencias son pequeñas y clínicamente poco relevantes y no coinciden con la ausencia de diferencias que muestra, en general, la literatura científica previa sobre el BDI-FS, pero, en cambio, son coherentes con los factores de riesgo demográficos encontrados en los estudios epidemiológicos sobre la depresión en AP en España.

10. Se han obtenido para el BDI-FS **baremos** generales y diferenciales en función del sexo para interpretar las puntuaciones de los pacientes españoles de AP. Esta es otra aportación novedosa y eminentemente práctica de la presente tesis doctoral tanto a nivel nacional como internacional, ya que su uso facilita de una manera sencilla el posicionamiento de un pacientes respecto a su grupo normativo y, por tanto, permite precisar la información sobre la gravedad de la sintomatología depresiva que ofrecen las categorías de gravedad en función de rangos de puntuaciones propuestas por Beck et al. (2000).

11. Aunque no era un objetivo específico de la presente investigación, el BDI-FS presenta **mejores índices de validez discriminante y de validez factorial** en pacientes de AP que la subescala de depresión de la HAD y la SDS. Esta información debería servir también como elemento de referencia a la hora de que los profesionales e investigadores españoles elijan el instrumento de evaluación y cribado de la depresión más adecuado para sus objetivos.

En resumen, a partir de los resultados de la presente investigación empírica se está en disposición de responder a diferentes preguntas teóricas y

prácticas que surgen en el ámbito de la evaluación y cribado de la depresión en AP, entre las cuales cabe reseñar dos:

- 1. ¿Mejora el uso de instrumentos de cribado la capacidad de detección de trastornos depresivos del médico de AP?** Sí. Tanto el BDI-FS como la subescala de depresión de la HAD superan al juicio clínico del médico de AP en la mayoría de los índices de rendimiento diagnóstico estudiados. Estos resultados, pues, reafirman el uso del cribado de depresión en AP.
- 2. ¿Ofrecen todos los instrumentos de cribado los mismos resultados para la detección de trastornos depresivos en AP?** No. Al menos en pacientes españoles de AP, el BDI-FS presenta índices de rendimiento diagnóstico mejores que los de la subescala de depresión de la HAD, instrumento que goza de una gran reputación y uso en España con todo tipo de pacientes médicos. Es más, parece que el BDI-FS también presenta en pacientes españoles de AP mejores índices de validez factorial y de validez discriminante que la subescala de depresión de la HAD y la SDS de Zung, por lo que a la hora de discriminar entre ansiedad y depresión, el BDI-FS podría ser una alternativa más válida que esos otros instrumentos.

Por supuesto, estas conclusiones deben valorarse a la luz de las limitaciones de la presente investigación, las cuales marcan además objetivos y prioridades para la investigación futura. Por ejemplo, sería deseable que futuros estudios pudieran examinar las propiedades psicométricas del BDI-FS a partir de una muestra de pacientes de AP más numerosa y seleccionada mediante procedimientos de muestreo aleatorio. También sería interesante

para posteriores estudios que los investigadores que realicen las evaluaciones no conozcan los objetivos de la investigación para así evitar posibles sesgos (no obstante, es importante señalar que en el primer estudio la especialista que realizaba el diagnóstico mediante la SCID-I no conocía los datos clínicos procedentes del médico de AP hasta que no terminaba la entrevista diagnóstica).

En cualquier caso, dado que los resultados de la presente tesis doctoral son positivos y alentadores, parece oportuno continuar con esta línea de investigación sobre la evaluación y cribado de la depresión en AP, especialmente a partir del BDI-FS, ya que la misma podría generar nuevos conocimientos ciertamente relevantes para la práctica clínica. En concreto, por ejemplo, el análisis de la utilización del BDI-FS en la práctica clínica cotidiana podría arrojar aportaciones muy interesantes al analizar si, frente a un centro de salud que no utilice instrumentos de evaluación y cribado de la depresión, en un centro que utilice el BDI-FS:

- ¿Aumentará el número de personas con depresión detectadas?
- ¿Recibirán los pacientes con depresión un mejor tratamiento?

También sería interesante ampliar el estudio de las propiedades psicométricas del BDI-FS a otros tipos de pacientes médicos (p. ej., hospitalizados, geriátricos, adolescentes) y a otros objetivos de evaluación (p. ej., la monitorización de los tratamientos antidepresivos), puesto que, aunque los resultados de esta tesis doctoral auguran que en esas otras poblaciones de pacientes y para esos otros objetivos el BDI-FS también puede ser muy válido y útil, es deseable su confirmación empírica específica.

Como síntesis final, esta tesis doctoral ofrece a los médicos de AP un instrumento con garantías para detectar trastornos depresivos entre sus pacientes y para cuantificar de manera fiable y válida la gravedad de la sintomatología depresiva que presenten.

Referencias bibliográficas.

- Agüera, L., Henríquez, R. y Llorente, D. (2001). Introducción conceptual. En Agüera L.(Ed). *Manejo de la depresión en Atención Primaria*. Madrid: IM&C, SA.
- Agüera, L., Failde, I., Cervilla, J.A., Díaz Fernández, P. y Mico J.A. (2010). Medically unexplained pain complaints are associated with underlying unrecognized mood disorders in primary care. *Family Practice*,2, 11-17.
- American Psychiatric Association (2002). *DSM-IV TR: manual diagnóstico y estadístico de los trastornos mentales*. Masson SA. Barcelona.
- Ani, C., Bazargan, M., Hindman, D., Farooq, M.A., Akhanjee, L., Yemofio F., Baker R. y Rodriguez M. (2008). Depression symptomatology and diagnosis: Discordance between patients and physicians in primary care settings. *BioMed Central* 9,1. Disponible en: <http://biomedcentral.com/1471-2296/9/1>.
- Aragónés, E., Gutierrez, M., Pino, C., Lucena, J., Virgili, C. y Garreta, I. (2001). Prevalencia y características de la depresión mayor y la distimia en Atención Primaria. *Aten Primaria*, 27; (9), 623-628.
- Aragónés, E., Masdéu, M., Cando, G. y Coll, B. (2001). Validez diagnóstica de la Self Rating Depression Scale de Zung en pacientes de Atención Primaria. *Actas Españolas de Psiquiatría*, 29 (5), 310-316.
- Aragónés, E., Gutierrez, M.A., Piñol, J.L., Mélich, N., Folch, S. y Labad, A. (2003). Prevalencia, expresión e impacto de los trastornos depresivos en atención primaria. *Atención Primaria*, 31, 120-125.
- Aragónès, E., Piñol, J.L; Labad, A. y Masdéu, R.M. (2004). Prevalence and determinants of depressive disorders in primary care practice in Spain. *International Journal of Psychiatry in Medicine*; 34, (1), 21-35.
- Aragónès E. (2005). Estudio de los trastornos depresivos en la atención primaria de salud. (Resumen de tesis doctoral). *Revista de Psiquiatría de la Facultad de Medicina de Barcelona*, 32, (1), 30-37.
- Aragónés, E., Piñol, L y Labad, A. (2006). Comorbilidad de la depresión mayor con otros trastornos mentales comunes en pacientes de Atención Primaria. *Atención Primaria*. 41, (10), 545-551.
- Arbesú, J.A., González, V.M.y Torquero, F. (2006). *Libro blanco de la depresión y ansiedad*. Madrid: Ed. IM&C, SA.

- Archer, R.P, Maruish, M, Imhof, E.A. y Piotrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology Research and Practice*, 22, 247–52.
- Arnau, R., Meagher, M., Norris, M., y Bramson, R. (2001). Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. *Health Psychology*, 20, 119-120.
- Arroll, B., Khin, N. y Kerse, N. (2003). Screening for depression in primary care with two verbally asked questions: cross sectional study. *British Medical Journal*, 32, 1144- 6.
- Ayuso-Mateos, J.L y Vazquez Barquero, J.L.(2001). Depressive disorder in Europe: prevalence figure from the ODIN study. *British Journal of Psychiatry*, 179, 308-316.
- Baca, E., Sainz, J., Agüera, L., Caballero, L., Fernández Liria, A., Ramos, J., Gil, A., Madrigal, M. y Porras, A. (1999). Validación de la versión española del PRIME-MD: Un procedimiento para el diagnóstico de los trastornos mentales en Atención Primaria. *Actas Españolas de Psiquiatría* 27 (6), 375-383.
- Bagby, MR, Ryder, A, Schuller, D y Marshall, M.(2004). The Hamilton Depression Rating Scale: Has the Gold Standard Become a Lead Weight? *The American Journal of Psychiatry* 161, 2163-2177.

Disponible en: <http://ajp.psychiatryonline.org/cgi/content/abstract/161/12/2163>
- Barkow, K., Maier, W., Bedirhan, T., Gäsicke, M., Wittchen, H-U y Heun, R. (2003). Risk factors for depression at 12-month follow-up in adult primary health care patients with major depression: an international prospective study. *Journal of Affective Disorders* 76, 157–169.
- Baumeister, H. y Harter, M. (2007). Prevalence of mental disorders based on general population surveys. *Social Psychiatry and Psychiatric Epidemiology* 42, 537–546.
- Bech, P., Malt, U.F., Dencker, S. J. Ahlfors, U.G. Elgen, K., Lewander, T., Lundell, A, Simpson, G.M. y Lingjmd, O.(eds).(1993). Scales for assesment of diagnosis and severity of mental disorders. *Acta Psychiatrica Scandinavica*; 87 (372), 7-81.

- Beck, A. T., Ward, C.H., Mendelson, M., Mock, J., y Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Beck, A. T. (1963). Thinking and depression. *Archives of general Psychiatry*, 9, 324-333.
- Beck, A.T., y Beck, R. W. (1972). Screening depressed patients in family practice rapid technic. *Postgraduate Medicine*, 52, 81-85.
- Beck, A.T. (1976). *Cognitive Therapy and the Emotional Disorders*. England: Oxford International Universities Press.
- Beck, A.T., Rush, A.J., Shaw, B.F. y Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press (Trad. esp. *Terapia Cognitiva de la Depresión*. (1983). Bilbao: Desclée de Brower,.
- Beck, A. T., Epstein, N., Brown, G., y Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56, 893-897.
- Beck, A. T., Steer, R.A y Garbin, M.C. (1988). Psychometric properties of the Beck Depression Inventory: Twenty five years of evaluation. *Clinical Psychology Review*, 8, 77-100.
- Beck, A. T y Steer, R. (1990). *Beck Anxiety Inventory manual*. San Antonio, TX: The Psychological Corporation.
- Beck, A. T. y Steer, R. (1993). *Beck Anxiety Inventory manual*. San Antonio, TX: The Psychological Corporation.
- Beck, A.T., Steer, R.A. y Brown, G.K. (1996). *Manual for the Beck Depression Inventory -II*. San Antonio, TX: The Psychological Corporation.
- Beck, A. T., Steer, R. A., Ball, R., Ciervo, C. A., y Kabat, M. (1997). Use of the Beck Anxiety and Beck Depression Inventories for primary care with Medical Outpatients. *Assessment*, 4, 211-219.
- Beck, A., Guth, D., Steer, J. y Ball, R. (1997). Screening for major depression disorders in medical inpatients with the Beck depression inventory for primary care. *Behaviour, Research and Therapy*. 35, (8), 785-791.

- Beck, A., Steer, R. y Brown, G. (2000). *BDI Fast Screen for medical patients*. San Antonio, TX: The Psychological Corporation.
- Bellón, J.A., Moreno Küstner, B., Torres, F., Montón Franco, C., Gil de Gómez, M.J., Sánchez Celaya, M., Díaz-Barreiros, M.A., Vicens, C., Luna, J.D., Cervilla, J., Gutierrez, B., Martinez Cañavate, M.T., Oliván Blázquez, B., Vázquez Medrano, A., Sánchez Artiaga, M.S., March, S., Motrico, E., Ruiz García, V.M., Brangier, R., Muñoz-García, M.M., Nazareth, I y King, M. for the predict group.(2008). Predicting the onset and persistence of episodes of depression in primary health care. The predictD-Spain study: Methodology. *BioMed Central Public Health* 8, 256.
- Benedict, R. H., Fishman, I., McClellan, M.M., Bakshi, R y Weinstock- Guttman, B. (2003). Validity of the Beck Depression Inventory-Fast Screen in multiple sclerosis. *Multiple Sclerosis Journal*, 9 (4), 393-6.
- Benitez, J.M., Casquero, R., Sánchez, F., Villalba, E. y Zamorano, E. (2009). *Depresión*. En SEMERGEN DoC. (2ª Edición). Depresión y Ansiedad (p. 9-23). Madrid. Edicomplet.
- Bennett, D. S., Ambrosini, P. J., Bianchi, M., Barnett, D., Metz, C. y Rabinovich, H. (1997). Relationship of Beck Depression Inventory factors to depression among adolescents. *Journal of Affective Disorders*, 45, 127-134.
- Berardi, D., Leggieri, G., Berti Ceroni, G., Rucci, P., Pezzoli, A., Paltrinieri, E. Grazian, N., y Ferrari, G. (2002). Depression in primary care. A nationwide epidemiological survey. *Family Practice*, 19, 397-400.
- Binns, H.J., Lannier, D., Pace, W.D., Galliher, J.M., Ganiats, T.G., Grey, M., Ariza, A.J. y Williams, R. (2007). Describing primary care encounters: The primary care network survey and the national ambulatory medical care survey. *Annals of Family Medicine*, 5, 38-47.
- Bjelland, I., Dahl, A., Haug, T.T., y Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *Journal of Psychosomatic Research*, 52, 69-77.
- Bobes, J.G., Portilla, M.P., Bascaran, M.T , Saiz PA., y Bousoño M, (2002). *Banco de instrumentos básicos para la práctica de la psiquiatría clínica*. Barcelona: Psiquiatría editores (Ars Médica).

- Bobes, J., Bulbuena, A., Luque, A., Dal-Ré, R., Ballesteros, J. y Ibarra, N., (2003).
Evaluación psicométrica comparativa de las versiones en español de 6, 17, 21 ítems en la escala de valoración de Hamilton en la depresión. *Medicina Clínica (Barcelona)*, 120, (18), 693-670.
- Bohnanmon, R. W., Maljanian R., y Goethe, J. (2003). Screening for depression in clinical practice: reliability and validity of a 5-item subset off CES-D. *Perceptual & Motor Skills*, 97, 855-861.
- Bollen, K. A., y Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 100, 305-314.
- Brennan, C., Worrall-Davies, A., McMillan, D., Gilbody, S. y House A. (2010). The Hospital Anxiety and Depression Scale: A diagnostic meta-analysis case-finding ability. *Journal of Psychosomatic Research* 69; 371–378.
- Briggs, S. R, y Cheek, J. M. (1986). The role of factor analysis in the evaluation of personality scales. *Journal of Personality*, 54,106-148.
- Boarman, J. y Walters, P. (2009) Managing depresión in primary care: it´s not only what you do it´s the way that you do it. *British Journal of General Practice*. 59 (560), 211-2. discusión 212.
- Bumberry, W., Oliver, J.M., y McClure, J.N. (1978). Validation of the Beck Depression Inventory in a university population using psychiatric estimate as the criterion. *Journal of Consulting and Clinical Psychology*, 46, 150-155.
- Caballero, L. Aragonés, E., García-Campayo, J., Rodríguez-Artalejo, F., Ayuso-Mateos, J.L: Polavieja, P., Gómez –Utrero, E., Romera, I. y Gilaberte, I. (2008). Prevalence, Characteristics and Attribution of Somatic Symptoms in Spanish Patients With Major Depressive Disorder Seeking Primary Health Care. *Psychosomatics*, 49, (6), 520-529.
- Caballero L. y Garcia- Parajúa P. (2004). Tratamiento de la depresión en Atención Primaria. En Inmaculada Gilaberte (Ed.). *Nuevas Perspectivas en la depresión*. Madrid: Aula Médica Ediciones.

- Carulla, S., López Alberca, R. y González Sáiz, F. (2000). Guías para la descripción y la selección de instrumentos de evaluación en psiquiatría. En Bulbena A., Berrios G., Fernández de Larrinoa P. *Medición clínica en psiquiatría y psicología*. Barcelona: Ed. Masson.
- Chamorro García, L. (2004). *Guía de manejo de los trastornos mentales en Atención primaria*. Barcelona: Psiquiatría Editores.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284-290.
- Chinchilla, A. (2008). *La depresión y sus mascararas. Aspectos terapéuticos*. Madrid: Editorial médica panamericana.
- Caro, I. e Ibáñez, E. (1992) La Escala Hospitalaria de Ansiedad y Depresión: su utilidad práctica en psicología de la salud. *Boletín de Psicología*; 36, 43-69.
- Cepoiu, M., McCusker, J., Cole, M.G., Sewitch, M., Belzile, E.y Ciampi, A.(2008). Recognition of depression by non-psychiatric physicians— a systematic literature review and meta-analysis. *Journal of General Internal Medicine*, 23, 25–36.
- Chisholm, D., Amir, M., Fleck, M., Herrman, H., Lomachenkov, A., Lucas, R., Patrick, D., (2001).Longitudinal Investigation of depression outcomes (the LIDO study) in primary care in six countries: comparative assesment of local health systems and resources utilisation. *International Journal of methods in Psyquiatric Research*, 10; 59-71.
- Chochinov, H.M. Wilson, K.G. Enns, M. & Lander, S., (1997) “Are you depressed?” Screening for depression in the terminally ill. *American Journal of Psychiatry*, 154, 674-6.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2ª ed. Hillsdale, NJ: LEA.
- Cohen, L. y Manion, L. (1990). *Métodos de investigación educativa*. Madrid: Ed. La Muralla.
- Conde, V., y Esteban, T. (1975). Fiabilidad de la S.D.S. (Self-Rating Depresión Scale) de Zung. *Revista de Psicología General y Aplicada*, 30, 903-913

- Conde, V. (1976) La medida de la depresión. *La escala autoaplicada para la medida cuantitativa de la depresión de Zung*. Graficesa, Facultad de Medicina de la Facultad de Salamanca, 1967-1969.
- Conde, V., Escriba, J.A. e Izquierdo, J.A. (1970). Evaluación estadística y adaptación castellana de la Escala Autoaplicada para la Depresión de Zung. *Archivos de Neurobiología*. Parte I, XXXIII, 2, 185-206 y Parte II, XXXIII, B, 201-203.
- Conde, V y Esteban-Chamorro, T. (1973). Revisión crítica de dos adaptaciones castellanas de la "Self Rating Depresión Scale" (SDS) de Zung. *Archivos de Neurobiología*, 36, 375-392.
- Cronbach, L.G. (1990). *Essentials of psychological testing*. Nueva York: Harper & Row.
- Cuijpers, P., Straten van A., Schaik van A. y Andersson G. (2009). Psychological treatment in primary care: a meta-análisis. *British Journal of general Practice*, 59, (559), e51-e60.
- Curran, C., Knapp, M., McDaid, D., Tomasson, K., the MHEEN Group, (2007). Mental health and employment: an overview of patterns and policies across 17 MHEEN countries. *Journal of Mental Health* 16 (2), 195-210.
- De Las Cuevas Castresana, C., García-Estrada Pérez, A., y González De Rivera, J.L. (1995). Hospital Anxiety and Depression Scale y psicopatología afectiva. *Anales de Psiquiatría*, 11(4), 126-130.
- Demyttenaere, K., Verhaeghen, A., Dantchev, N., Grassi, L., Montejo, A.L., Perahia, D., MRCPsych., Quail, D., Reed, C. y Tylee, A., FRCGP, Bauer M. (2009). "Caseness" for Depression and Anxiety in a Depressed Outpatient Population: Symptomatic Outcome as a Function of Baseline Diagnostic Categories *Primary Care Companion Journal of Clinical Psychiatry*, 11(6), 307-315.
- Díaz Suarez, J. Martínez Borrondo, M., Bousoño García, M., Bobes, J. (2004). Valoración Clínica de los comportamientos suicidas. En Bobes, J., Saíz, P., García-Portilla, MP., Bascarán, MT. y Bousoño, M. *Comportamientos suicidas. Prevención y Tratamiento*. Barcelona: Ed. Ars Médica.

- Diez-Quevedo, C., Rangil, T., Sanchez-Planell, L., Kroenke K., y Spitzer, R. (2001).
Validation and Utility of the Patient Health Questionnaire in Diagnosing Mental Disorders in 1003 General Hospital Spanish Inpatients. *Psychosomatic Medicine*, 63, 679-686.
- Dowell, A. C., y Biran, L. A. (1990). Problems in using the hospital anxiety and depression scale for screening patients in general practice. *British Journal of General Practice*, 40, 27-28.
- Dowrick, C.F., Bellon, J.A.y Gómez, M.J. (2000) GP frequent attendance in Liverpool and Granada: the impact of depressive symptoms. *Br J Gen Pract*, 50, 361-365.
- Doyle B. (1990) Crisis management of the suicidal patient. En Blumenthal Sj, Kupfer DJ (eds.) *Suicide over the life cycle*. Washington: American Psychiatric Press, 381-423.
- Dozois, D., Dobson, K., y Anhberg, J. (1998). A psychometric evaluation of the Beck Depression Inventory II. *Psychological Assessment*, 10, 83-89.
- Dutton, G., Groethe, K. y Jones, G. (2004). Use of the Beck Depression Inventory II with African American primary care patients. *General Hospital Psychiatry*, 26, 437-442.
- Eguiluz, J. I. (Editor). (2001). *Introducción a la psicopatología*. Madrid: IM&C, S.A.
- Eaton, W., Shao, H., Nestadt, G. Lee, B.H., Bienvenue, O.J. y Zandi, P (2008) Population-based study of first onset and chronicity in major depressive disorder. *Archives of General Psychiatry* 65 (5), 513-520.
- Fernández-Ballesteros, R.(dir.). (2010) *Evaluación psicológica. Conceptos, métodos y estudio de casos*. Madrid: Ediciones Pirámide.
- Fernández-Alonso, M.C., Buitrago-Ramírez, F., Ciurana-Misol, R., Chocrón- Bentata, L., García-Campayo, J., Montón-Franco, C y Tizón-García, J. Grupo de trabajo de Salud Mental del Programa de Actividades Preventivas y de Promoción de la Salud (PAPPS).(2009). *Guía de Salud Mental en Atención Primaria*. Barcelona: Sociedad Española de Medicina de Familia y Comunitaria (semFYC).

Disponible

en:

<http://www.papps.org/upload/file/02%20PAPPS%20ACTUALIZACION%202009.pdf>

- Fernández, L.J. (1999). *Aspectos básicos de Salud Mental en Atención Primaria*.
Madrid: Editorial Trotta.
- Feightner, J. W., y Worrall, G. (1990). Early detection of depression by primary care physicians. *Canadian Medical Association Journal*, 142, 1215-1220.
- First, M. B., Spitzer, R. L., Gibbon, M. y Williams, J. B. W. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders—Research Version-Patient Edition*. New York: Biometrics Research.
- First, M., Spitzer, R., Gibbon, M. y Williams, J., (1996). *Guía de usuario para la entrevista clínica estructurada para los trastornos del eje I del DSM IV. Versión clínica*. Washington : American Psychiatric Press, Inc. Masson.
- First, M., Spitzer, R., Gibbon y Williams, J. (1999). *Entrevista Clínica Estructurada para los trastornos del Eje I del DSM-IV (versión clínica) SCID-I*. Barcelona: Masson.
- Furlanetto, L. M., Mendlowicz, M. V. y Romildo Bueno, J. (2005). The validity of the Beck Depression Inventory-Short Form as a screening and diagnostic instrument for moderate and severe depression in medical inpatients. *Journal of Affective Disorders*, 86, 87-91.
- Fischer, J. E., Bachmann, L. M., y Jaeschke, R. (2003). A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Medicine*, 29, 1043-1051.
- Gabarrón, H., Vidal, R., Haro, A., Boix, S., Jover, B., y Arenas, P. (2002). Prevalencia y detección de los trastornos depresivos en atención primaria. *Atención Primaria*, 29,(6), 329-337.
- Gabilondo, A., Rojas-Farreras, S., Vilagut, S., Haro, J.M., Fernández, A., Pinto- Meza, A. y Alonso, J. (2010). Epidemiology of major depressive disorder in a southern European Country: Results from the ESEMED-Spain project. *Journal of Affective Disorders* 120, 76-85.
- García-Campayo, J., Sanz, C., y Jimeno, A. (1993). Interconsulta en Atención Primaria /psiquiatría: utilización de la consulta única. *Atención Primaria*, 13, 52-55.

- García-Campayo, J., Ayuso-Mateos, J.L. , Caballero, L., Romera, I., Aragonés,E., Rodriguez Artalejo, F., Quail, D., y Gilaberte, I. (2008). Relationship of somatic symptoms with depression severity, quality of life, and health resources utilization in patients with major depressive disorder seeking primary health care in Spain. *Prim Care companion Journal of Clinical Psychiatry*, 10, 355-362.
- García-Testal, A., Sancho Blasco, F., Julve Pardo, R., Puche Pinazo, E., Rabanaque, B. y Mallén, G. (1998). Estudio de las derivaciones de atención primaria a salud mental: ¿qué coincidencia existe entre los motivos de derivación y el diagnóstico del especialista? *Atención Primaria*, 22, 233-238.
- Gibbons, R., Clark, D. y Kupfer, D. (2002). Exactly what does the Hamilton depression rating scale measure? *Journal of Psychiatric Research*, 27, (3), 259-273.
- Gilbody, S. H., House, A. P., y Sheldon, T. A. (2001). Routinely administered questionnaires for depression and anxiety: a systematic review. *British Medical Journal*, 322, 406-409.
- Gilbody, S., Richards, D., Brealey, S. y Hewitt, C.(2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *Journal of General Internal Medicine*, 22, 1596–602.
- Golden J.,Convoy R. y O'Dwyer A. (2007). Reliability and validity of the Hospital Anxiety and Depression Scale and the Beck Depression Inventory (Full and Fast Screen scales) in detecting depression in persons with hepatitis C. *Journal of Affective Disorder*, 100, 265-269.
- Goldberg, D., Bridges, K. y Duncan-Jones, P., Grayson, D. (1988) Detecting anxiety and depression in general medical settings. *British Medical Journal*; 97, 897-899.
- Goldberg, D., Bridges, K. y Duncan-Jones, P., Grayson, D. (1989). Detección de la ansiedad y la depresión en el marco de la medicina general. *British Medical Journal (ed. Esp.)*;4(2), 49-53.
- Goldberg, D. (2006). The “NICE Guideline” on the treatment of depression. *Epidemiologia e Psiquiatria Sociale*. 15 (1), 11-15.

- Goldberg D. (2006). *NICE clinical guideline 91 Depression in adults with a chronic physical health problem Treatment and management* .Developed by the National Collaborating Centre for Mental Health .
- Goldberg, D., Kendler, K., Sirovatka, P. y Regier, D. (2011). *Temas de investigación: DSM-IV. Depresión y trastorno de ansiedad generalizada*. Madrid: Editorial Médica Panamericana.
- Gray, J.(2004). New Concepts in screening. *British Journal of General Practice*, 54, 292-298.
- Greenberg, P.E y Kessler, R.C.(2003). The economic burden of depression in the United States: how it did change between 1990 and 2000? *Journal of Clinical Psychiatry*, 64, 1465-75.
- Grupo de trabajo de Salud Mental del Programa de Actividades Preventivas y de Promoción de la Salud (PAPPS). Guía de Salud Mental en Atención Primaria. Barcelona. Sociedad Española de Medicina de Familia y Comunitaria (semFYC),. (2001) Disponible en:
<http://www.papps.org/upload/file/publicaciones/guiaRevisada2008-salud-mental.pdf>
- Grupo de Trabajo de la Guía de Práctica Clínica para el Manejo de Pacientes con Trastornos de Ansiedad en Atención Primaria (2008). *Guía de Práctica Clínica para el Manejo de Pacientes con Trastornos de Ansiedad en Atención Primaria*. Madrid: Plan Nacional para el SNS del MSC. Unidad de Evaluación de Tecnologías Sanitarias. Agencia Laín Entralgo. Comunidad de Madrid. Guías de Práctica Clínica en el SNS: UETS Nº 2006/10.
- Hall, A., A'Hern, R., y Fallowfield, L. (1999). Are we using appropriate self-report questionnaires for detecting anxiety and depression in women with early breast cancer? *European Journal of Cancer*, 35, 79-85.
- Henkel, V., Megl, R., Coyne, J. C., Kohnen, R., Moller, H.J. y Hegerl, U. Screening for depression in PC: will one or two items suffice? (2004). *European Archives of Psychiatry & Clinical Neurosciences*, 254, (4), 215-223.

- Hatzenbuehler, L. C., Parpal, M. y Matthews, L. (1983). Classifying college students as depressed or nondepressed using the Beck Depression Inventory: an empirical analysis. *Journal of Consulting and Clinical Psychology*, 51(3), 360-366.
- Healey, A.K., Kneebone, I.I: Carroll, M. y Anderson ,S.J. (2008). A preliminary investigation of the reliability and validity of the Brief Assessment Schedule Depression Cards and the Beck Depression Inventory-Fast Screen to screen for depression in older stroke survivors. *International Journal of Geriatric Psychiatry*, 23, 531–536.
- Herrman, S.M. y McKenna, S.P. (1992). The QLDS: a scale for the measurement of quality of life in depression. *Health Policy*, 22, 307-319.
- Herrero, J. y Gracia, E. (2007). Una medida breve de la sintomatología depresiva (CESD-7). *Salud Mental*, 30, 5.
- Herrero, M.J, Blanch, J., Peri, J.M., De Pablo, J, Pintor, L. y Balbuena, L. (2003). A validation study of the Hospital anxiety and depression scale (HADS) in a Spanish population. *General Hospital Psychiatry*, 25, 277-283.
- Hickie, I., Davenport, T., y Ricci, C. (2002). Screening for depression in general practice and related medical settings. *The Medical Journal of Australia*, 177, S111-S116.
- Johnson, G., Burvill, P.W., Anderson, C.S., Jamrozik, K., Stewart-Wynne, E.G., Chakera, T.M.H.S (1995). Screening instruments for depression and anxiety following stroke: experience in the Perth community stroke study. *Acta Psychiatrica Scandinavica*, 91, 252-257.
- Joiner, Jr. T.E., Walker, R.L., Pettit, J.W., Perez, M. y Cukrowicz, K.C. (2005). Evidence based assessment of depression in adults. *Psychological Assessment*, 17, 267–77.
- Karapampa K., Borgstrom F. y Jonsson B. (2011). Economic burden of depression on society. *Medicographia*, 33, (2), 163-168.

- Karen, B. Grothe, K.B., Dutton, G.R., Jones G.J., Bodenlos J., Ancona M. y Brantley P. (2005). Validation of the Beck Depression Inventory—II in a Low-Income African American Sample of Medical Outpatients. *Psychological Assessment* 17, (1), 110–114.
- Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L., e Ingram, R.E. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. *Cognitive Therapy and Research*, 11(3), 289-299.
- Kessler, R., Berglund, P., Demler, O. Jin, R., Koretz, D., Merikangas, K., Rush, J.,Walters, E.y Wang, P. (2003). The epidemiology of Major Depressive Disorder: Results from de national comorbidity survey replication (NCS-R). *The Journal of American Medical Association*, 289(23):3095-3105.
- Kessler, R., Chiu, WT., Demler, O. y Walters, E. (2005). Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry*, 62, 617-627.
Disponible en: www.archgenpsychiatry.com
- King, M., Nazareth, I., Levi, G., Walker, C., Morris, R., Weich, S., Bellon- Saameño, J.A., Moreno, B., Svab, I., Rotar, D., Rifel, j., Maarroos, H.I., Alioja, A., Kalda, R.,Neeleman, J., Geerlings, M.I., Caldas de Almeida, M., Coreea, B. y Torres-González, F. (2008). Prevalence of common mental disorder in general practice attendees across Europe.*The british journal of Psychiatry* 192, 362-367.
- Kohn, R., Levav I., Caldas de Almeida, J.M., Vicente, B., Andrade, L., Caraveo-Anduaga, J.J., Saxena, S.y Saraceno, B. (2005;) Los trastornos mentales en América Latina y el Caribe: asunto prioritario para la salud pública. *Revista Panamericana de Salud Publica*. 18(4/5), 229–40.
- Krefetz, D.G, Steer, R.A, Jermyn, R.T. y Condoluci, D.V. (2004). Screening HIV infected patients with chronic pain for anxiety and mood disorders with the beck anxiety and depression inventory fast screens for medical settings. *Journal of Clinical Psychology Medical Settings* 11, :283–9.
- Kriston, L. y Wolff, A., (2011) Not as golden as standards should be Interpretation of the Hamilton RatingScale for Depression Journal of Affective Disorders 128 175–177

- Kroenke, K., Spitzer, R.L. y Williams, J.B., (2001). The PHQ-9: Validity of a brief depression severity measure, *Journal of General Internal Medicine.*, 16 (9), 606-13.
- Kroenke, K. y Spitzer, R.L., (2002) The PHQ-9: A new depression and diagnostic severity measure. *Psychiatric Annals*, 32, 509-21.
- Kroenke, K., Strine, T., Spitzer, R.L., Williams J.B., Berry J.T. y Mokhad A.H. (2009) The PHQ-8 as a measure of current depression in the general population. *Journal of affective disorders* 114, 163-173.
- Krueger, R.F., Caspi, A., Moffitt, T.E., Silva, P.A. (1998). The structure and stability of common mental disorders (DSMIII-R): A longitudinal/epidemiological study. *Journal of Abnormal Psychology*, 107, 216-227.
- Lake, C. R. (2008) How Academic Psychiatry Can Better Prepare Students for Their Future Patients: Part I: The Failure to Recognize Depression and Risk for Suicide in Primary Care; Problem Identification, Responsibility, and Solutions. *Behavioral Medicine*, 34, (3), 95-100.
- Lake, C.R. (2008). How Academic Psychiatry Can Better Prepare Students for Their Future Patients: Part II: A Course in Ultra-Brief Initial Diagnostic Screening Suitable for Future Primary Care Physicians. *Behavioral Medicine*, 34: (3), 101-118.
- Lam, N., Contreras, H. Mor, E., Gil, E., Espinoza, E., y Córdova, G. Estudio de la validez interna y concurrente del inventario de depresión de Beck para atención primaria (BDI-PC) en población gestante de Lima Metropolitana, mayo – junio de 2008. (2009). *Revista peruana. Epidemiología*, 13,(2), 28-32.
- Lasa, L., Ayuso-Mateos, J. L., Vázquez-Barquero, J. L., Díez-Manrique, F. J., y Dowrick, C. F. (2000). The use of the Beck Depression Inventory to screen for depression in the general population: a preliminary analysis. *Journal of Affective Disorders*, 57, 261–265.

- Linárez, R. (2005). *El reconocimiento de la depresión en atención primaria*.
(Universidad: de Málaga) Cota: TA WM171 L55 2004 (434) Enlace a resumen de tesis:
http://www.kriptia.com/PSICOLOGIA/PSICOLOGIA_EXPERIMENTAL/ANALISIS_EXPERIMENTAL_DE_LA_CONDUCTA/1#118531.
- Lepine, J.P., Gastpar, M., Mendlewicz, J. y Tylee, A. (1997) Depression in the community: the first pan-European study DEPRES (Depression Research in European Society). *International Clinical Psychopharmacology*, 12, (1), 19-29.
- Leng, J., Changrani, J., Tseng, C.H. y Gany, F., (2010). Detection of Depression with Different Interpreting Methods Among Chinese and Latino Primary Care Patients. A Randomized Controlled Trial. *Journal of Immigrant Minority Health* 12, 234-241.
- Levine, T. R., y Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, 28, 612-625.
- Lobbestael, J., Leurgans, M., y Arntz, A. (2011). Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). *Clinical Psychology & Psychotherapy*, 18, 75-79.
- Lopez, A., Mathers, Colin, D., Ezzati, M., Jamison, D. y Murray, J.L. (2006). Global and regional burden of disease and risk factors, 2001: Systematic analysis of population health data. *Lancet*, 367, 1747-1757.
- López-Pina, J.A., Sánchez-Meca, J. y Rosa-Alcázar, A.I. (2009). The Hamilton Rating Scale for Depression: A meta-analytic reliability generalization study. *International Journal of Clinical and Health Psychology*. 9, (1), 143-159.
- López-Roig, S., Terol, M. C., Pastor Mira, M. Á., Neipp, M. C., Massutí, B., Rodríguez Marín, J. (2000). Ansiedad y depresión. Validación de la Escala HAD en pacientes oncológicos. *Revista de Psicología de la Salud*, 12(2), 127-155.

- Love, A.W., Grabsch, B., Clarke, D.M., Bloch, S. y Kissane, D. W. (2004). Screening for depression in women with metastatic breast cancer: a comparison of the Beck Depression Inventory Short Form and the Hospital Anxiety and Depression Scale. Australian and New Zealand. *Journal of Psychiatry*, 38, 526-531.
- Lowe, B., Unutzer, J., Callahan, C.M., Perkins, A.J. y Kroenke, K. (2004). Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. *Medical Care Journal*, 42, 1194–201.
- Mackinnon, A., Christensen, H., Jorm, A.F., Henderson, A. S., Scott, R. y Korten, A.E. (1994) A latent trait analysis of an inventory designed to detect symptoms of anxiety and depression using an elderly community sample. *Psychological Medicine Journal*, 24, 977-986.
- MacMillan, H.L., Patterson, C.J.S., y Whathen, C.N. (2005). Screening for depression in Primary Care: recommendation statements from the Canadian Task Force on Preventive Health. *Canadian Medical Association Journal* 4, 172, (1), 33-35.
- Magruder K.M. y Yeager D.E. Screening for depression in primary care: Can it become more efficient? En Mitchel A.J. y Coyne J.C.(2010). *Screening for depression in primary clinical practice. An evidence-based guide*. New York. Oxford University Press.
- Mc Grady, A., Lynch, D.J, Nagel, R.W. y Tamburrino, M. (2010). Coherence between physician diagnosis and patient self report of anxiety and depression in primary care. *The journal of nervous and mental disease*, 198, 6, 420-424.
- Magán, I., Sanz, J., y García-Vera, M. P. (2008). Psychometric properties of Spanish version of the Beck Anxiety Inventory (BAI) in general population. *The Spanish Journal of Psychology*, 11, 624-640.
- Marsh, H. W. (1996). Positive and negative global self-esteem: a substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, 70, 810-819.
- Marquis, P.(1998). Strategies for interpreting Quality of Life Questionnaires. *Quality of Life Newsletter*, 3-5.

- Martín-Águeda, B., López-Muñoz, F., Álamo, C., García-García, P., Silva, A., y Rubio, G. (2005). Situación actual de la depresión en España en el ámbito de la atención primaria: aspectos asistenciales, clínicos y terapéuticos. *APySAM. Atención Primaria y Salud Mental*, 8, 22-31.
- Martín-Águeda, B., López-Muñoz, F., Silva, A., García-García, P., Rubio, G. y Álamo, C. (2007) Aspectos diferenciales del manejo de la depresión en España entre la atención psiquiátrica especializada y la atención primaria. *Actas Españolas de Psiquiatría* 35, (2), 89-98.
- Mathers, CD y Loncar, D(2006) Projections of global mortality and burden of disease from 2002 to 2030. *Public Library of Science Medicine*, 3, (11): e442.
- MedCalc para Windows XP, versión 11.6.0.0. .*MedCalc Software*, Mariakerke, Bélgica. Disponible en <http://www.medcalc.org/>
- Mitchell, AJ y Coyne, JC. (2007) Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *British Journal of General Practice*, 57, 144–51.
- Mitchell, A. J., Vaze, A., y Rao, S. (2009) Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* , 374, 609–19.
- Montgomery, S.(2008). Core symptoms of depresión in clinical practice. *Medicographia*. 30, 52-54.
- Montón. C., Pérez-Echevarría, M.J., Campos, R., García Campayo, J. y Lobo, A. (1993) Escalas de ansiedad y depresión de Goldberg: una guía de entrevista eficaz para la detección del malestar psíquico. *Atención Primaria*,; 12, 345-349.
- Monsalve, V., Soriano, J., e Ibáñez, E.(2001). Fiabilidad y estructura factorial de la escala de ansiedad y depresión hospitalaria (HAD) en una muestra de pacientes con dolor crónico. *Boletín de Psicología*, 72, 7-22.
- Morriss, R., Leese, M., Chatwin, J. y Baldwin, D, (2008). Inter-rater reliability of the Hamilton Depression Rating Scale as a diagnostic and outcome measure of depression in primary care. *Journal of Affective Disorder*. 111(2-3), 204-213.

- Mulrow, C.D., Williams, J.W.Jr., Gerety, M.B., Ramirez, G., Montiel, O.M., y Kerber, C.(1995). Case-finding instruments for depression in primary care settings. *Annals of Internal Medicine*, 122, 913-921
- Muñiz, J. y Fernández-Hermida, J.R. (2000). La utilización de los tests en España. *Papeles del Psicólogo*, 76, 41-49.
- Muñiz, J., Fernández-Hermida, J.R., Fonseca-Pedrero, E., Campillo-Álvarez, A. y Peña-Suárez, E., (2011). Evaluación de los test editados en España. *Papeles del Psicólogo*, 32, (2), 113-128.
- Muñoz, P. (2000). Test de cribado en la práctica clínica. En Bulbena A., Berrios G., Fernández de Larrinoa P. (Ed). *Medición clínica en psiquiatría y psicología*. (p. 35-54). Barcelona: Masson.
- Murphy, GE: The psysician´s role in suicide prevention. En Roy A. (eds.) *Suicide*. (1986). (p. 171-179). Baltimore: Williams and Wilkins.
- Murray, C.J.L. y López, A.D. (Eds.). (2001). *The global burden of disease 2000 project: aims, method and data sources*. Geneva:World Health Organization (GPE discussion paper n.0 36), a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Cambridge, MA: Harvard University Press.
- National Screening Committee UK. (2003). *Criteria for appraising the viability, effectiveness and appropriateness of a screening programme*. London: National Screening Committee, UK; Visitado en: <http://www.screening.nhs.uk/screening>
- Neal, RM y Basldwin, RC(1994). Screening for anxiety and depression in elderly medical outpatients. *Age Ageing*; 23, 461- 464.
- NICE (2009). *Depression in adults with a chronic physical health problem.Treatment and Management*. London. National Institute for Health and Clinical Excellence. Disponible en www.nice.org.uk/CG91
- Nunnally, J. C., y Bernstein, I. J. (1995). *Teoría psicométrica*. México: McGraw Hill.
- O´Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallelanalysis and Velicers MAP test. *Behavior Research Methods, Instruments, and Computers*, 32, 396-402.

- Oliver, J.M., y Burckham, R. (1979). Depression in university students: duration, relation to calendar time, prevalence, and demographic correlates. *Journal of Abnormal Psychology*, 88, 667-670.
- Olivera, J., Benabarre S., Lorente, T., Rodríguez, M., Pelegrín, C., Calvo, J.M. y Leris, J.M., Idáñez, D., Arna, S. (2008) *International Journal of Geriatric Psychiatry*; 23, 915-921.
- Orive, M., Padierna, J.A., Quintana, J.M., Las Hayas, C., Vrotsou, K. y Aguirre, V. (2010). Detecting depression in medically ill patients: Comparative accuracy of four screening questionnaires and physicians' diagnoses in Spanish population. *Journal of Psychosomatic Research*, 69, 399–406.
- O'Rourke, N. (2004). Reliability generalization of responses by care providers to the Center for Epidemiologic Studies-Depression Scale. *Educational and Psychological Measurement*, 64, 973-990.
- Osman, A., Kopper, B., Barrios, F., Gutierrez, P., y Bagge, C. (2004). Reliability and validity of the Beck Depression Inventory II with adolescent psychiatric inpatients. *Psychological Assessment*, 16, 120-132.
- Palmer, S. y Coyne, J.C. (2003) Screening for depression in medical care: Pitfalls, alternatives, and revised priorities. *Journal of Psychosomatic Research*. 54, 279–287.
- Penley, J., Wiebe, J.S. y Nwosu, A. (2003). Psychometric Properties of the Spanish Beck Depression Inventory–II in a Medical Sample. *Psychological Assessment*, 15, (4), 569–577.
- Piacentini, J., Roper, M., Jensen, P., Lucas, C., Fisher, P., Bird, H., Bourdon, K., Schwab-Stone, M., Rubio-Stipec, M., Davies, M., y Dulcan, M. (1999). Informant-based determinants of symptom attenuation in structured child psychiatric interviews. *Journal of Abnormal Child Psychology*, 27(6), 417- 428.
- Pinto-Meza, A, Serrano-Blanco, A, Penarrubia, MT, Blanco, E. y Haro, JM.(2005) Assessing depression in primary care with the PHQ-9: can it be carried out over the telephone? *Journal of General Internal Medicine*, 20, 738–42.

- Pio de Almeida, M., Simon, G., Herman, H., Bushnell, D., Martin, M. y Patrick, D.
(2005). Major depression and its correlates in primary care settings in six countries: 9 month follow-up study. *The British Journal of Psychiatry*, 186, 41-47.
- Piotrowsky, C. (1996). Use of the Beck Depression Inventory in clinical practice.
Psychological Reports, 79, 873-874.
- Piotrowski, C., y Gallant, N. (1999). Research use of clinical measures for anxiety in the recent psychological literature. *Journal of Instructional Psychology*, 36(1), 84-86.
- Plan de salud departament de salut de la generalitat catalana. Disponible en:
<http://www.gencat.cat/salut/depsalut/pdf/2007lb16.pdf>
- Plumb, M. M. y Holland, J. (1977). Comparative studies of psychological function in patients with advanced cancer, I: Self-reported depressive symptoms. *Psychosomatic Medicine*, 39, 264-276.
- Poole, H, Bramwell, R, Murphy, P. (2009) The utility of the Beck Depression Inventory Fast Screen (BDI-FS) in a pain clinic population. *European Journal of Pain*, 13, 865-869.
- Prieto, G. y Muñoz, J.(2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65-71.
- Pygnone, M. P., Bradley N., Gaynes, B. N., Rushton, J. L., Burchell, C. M., Orleans, C. T., Mulrow, C. D. y Lohr, K. (2002). Screening for depression in adults: a summary of the evidence for de U.S. *Preventive Services Task Force. Annals of Internal Medicine*, 136, 765-776.
- Quintana, JM, Padierna A, Esteban, C, Arostegui, I, Bilbao, A, y Ruiz, I. (2003). Evaluation of the psychometric characteristics of the Spanish Version of the Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*, 107, 216-221.
- Ramos-Brieva, JA y Cordero Villafafila, A. (1986). Validación de la versión castellana de la Escala de Hamilton para la Depresión. *Actas Luso Española de Neurología Psiquiátrica y Ciencias Afines*, 14, 324-334

- Retolaza, A. (2006) *Instrumentos de Criba en Atención Primaria*. Comunicación presentada en XXIII Congreso de la Asociación Española de Neuropsiquiatría. Bilbao. Mayo de 2006.
- Reuland, D., Cherrington, A., Watkins, G (2009). Diagnostic Accuracy of Spanish Language Depresión-Screening Instruments. *Annals of family medicine*, 7 (5). 305-309.
- Sadock, B. y Sadock, V. (2005). *Kaplan y Sadock's Pocket Handbook of clinical psychiatry*. Fourth Edition. Philadelphia. Lippincot Williams & Wilkins.
- Sánchez Pérez, J. A., López Cárdenas, A., (2005). Escalas diagnósticas y de evaluación que se utilizan en atención primaria para depresión y ansiedad. *Salud Global Año V (3)*, 33-36.
- Sanz, J. (1991). The specific traits of anxiety in the Anxiety Situations and Responses Inventory (ASRI): Construct validity and relationship to depression. *Evaluación Psicológica/ Psychological Assessment*, 7, 149-173.
- Sanz, J. y Vázquez, C. (1991). *Trastornos depresivos (II): Productos, operaciones, proposiciones y estructuras cognitivas*, en A. Belloch y E. Ibañez (compls). Manual de psicopatología, vol. 2. Valencia. Promolibro.
- Sanz, J. (1993). Distinguiendo ansiedad y depresión: Revisión de la hipótesis de la especificidad de contenido de Beck. *Anales de Psicología*, 9, 133-170.
- Sanz, J. (1992). *Procesos cognitivos en la ansiedad y en la depresión*. Tesis doctoral. Universidad Complutense de Madrid. Disponible en internet: <http://eprints.ucm.es/tesis/19911996/S/4/S4004801.pdf>
- Sanz, J. y Vázquez, C. (1995). *Trastornos del estado de ánimo: Teorías psicológicas*. En A. Belloch y B. Sandin y F. Ramos (Eds). Manual de psicopatología, vol. 2. Páginas 341-378. Madrid. McGraw Hill. ISBN:84-7986-000-6.
- Sanz, J., y Vázquez, C. (1998). Fiabilidad, validez y datos normativos del Inventario para la Depresión de Beck. *Psicothema*, 10, (2), 303-318

- Sanz, J., Navarro, M. E., y Vázquez, C. (2003). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 1. Propiedades psicométricas en estudiantes universitarios. *Análisis y Modificación de Conducta*, 29, 239-288.
- Sanz, J., y Navarro, M. E. (2003). Propiedades psicométricas de una versión española del Inventario de Ansiedad de Beck (BAI) en estudiantes universitarios. *Ansiedad y Estrés*, 9, (1) 59-84.
- Sanz, J., Perdigón, L. A., y Vázquez, C. (2003). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 2. Propiedades psicométricas en población general. *Clínica y Salud*, 14, 249-280.
- Sanz, J., García-Vera, M. P., Espinosa, R., Fortún, M., y Vázquez, C. (2005). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 3. Propiedades psicométricas en pacientes con trastornos psicológicos. *Clínica y Salud*, 16, 121-142.
- Sanz, J., García-Vera, M. P., Espinosa, R., y Fortún, M. (2005). Desarrollo y propiedades psicométricas de una versión breve española del Inventario para la Depresión de Beck-II (BDI-II). Comunicación presentada en el V Congreso Iberoamericano de Evaluación Psicológica, Buenos Aires (Argentina), 1-2 de julio de 2005.
- Sanz, J., y García-Vera, M. P. (2007). Análisis psicométrico de las versiones breves del «Inventario para la Depresión de Beck» de 1978 (BDI-IA). *Psicología Conductual*, 15, 191-214
- Sanz, J., y García-Vera, M. P. (2009). The Beck Depression Inventory-second edition (BDI-II): factor congruence and generalizability of its indexes of internal consistency. En E. Řehulka (Ed.), *School and health 21. General issues in health education* (p. 331-342). Brno: MSD.
- Sanz, J., y García-Vera, M. P. (2011). *Rendimiento diagnóstico y estructura factorial del Inventario para la Depresión de Beck—Segunda Edición (BDI-II) en pacientes españoles con trastornos psicológicos*. Manuscrito bajo revisión editorial.
- Sanz, J., García-Vera, M. P., y Fortún, M. (2011). El Inventario de Ansiedad de Beck (BAI): propiedades psicométricas de la versión española en pacientes con trastornos psicológicos. Manuscrito bajo revisión editorial.

- Sanz, J. (en prensa). Evaluación de los trastornos del estado de ánimo. En J. L. Graña Gómez, M. E. Peña Fernández y J. M. Andreu Rodríguez (Eds.), *Manual docente de psicología clínica legal y forense, vol. II* . Madrid: Servicio de Publicaciones de la Universidad Complutense de Madrid.
- Serrano-Blanco, A., Palao, J., Luciano, V., Pinto Meza, A., Lujan, L., Fernández, A., Bertsch, J., Mercader, M. y Haro J. (2010). Prevalencia, detección y adecuación del tratamiento de los trastornos mentales a la Atención Primaria de Salud de Cataluña . *Social Psychiatry and Psychiatric Epidemiology*, 45, 201-210.
- Sharp, L.K. y Lipsky MS. (2002) Screening for depression across the lifespan: a review of measures for use in primary care settings. *American Family Physician*; 66,1001–8.
- Scheinthal, S., Steer, R., Giffin, L. y Beck, A. Evaluating geriatric medical outpatients with the Beck Depression Inventory-FastScreen for Medical Patients. *Aging & Mental Health* 2001; 5(2): 143–148.
- Silverstone, P. H. (1994). Poor efficacy of the Hospital Anxiety and Depression scale in the diagnosis of major depressive disorder in both medical and psychiatric patients. *Journal of Psychosomatic Research*, 38, 441-450.
- Simon, G.E., Goldberg, D., Ustun, T.B., (1999) Outcomes of recognized and unrecognized depression in an international primary care study. *General Hospital Psychiatry*, 21, (2), 97-105.
- Spector, P. E., Van Katwyk, P. T., Brannick, M. T., y Chen, P. Y. (1997). When two factors don't reflect two constructs: how item characteristics can produce artifactual factors. *Journal of Management*, 23, 659-677.
- Spinoven, P. H., Ormel, J., Sloekers, P. P., et al. (1997). A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychological Medicine*, 27, 363-370.
- Soler, J., Pérez Sola, V., Puigdemot, D, Pérez-Blanco, J. y Figueres, M. (1997). Estudio de Validación del Center for Epidemiologic Studies –Depression (CES-D) en una población española de pacientes con trastornos afectivos. *Actas Luso-Españolas de Neurología, Psiquiatría y ciencias afines*, 25, (4), 243-249.

- Somoza, E., y Mossman, D. (1992). Comparing diagnostic tests using information theory: The INFO-ROC technique. *Journal of Neuropsychiatry and Clinical Neurosciences*, 4, 214-219.
- Soriano, J. y Salavert, L. (1996) The Hospital Anxiety and Depression Scale questionnaire in physical danger and evaluation situations. *Counseling Psychology Journal*, 9, 281– 96.
- Spitzer, RL, Kroenke, K. y Williams, JB. (1999) Validation and utility of a self report version of PRIME-MD: the PHQ primary care study. *Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. Journal of American Medicine Association*; 282, 1737–44.
- Scheinthal, SM, Steer, RA, Giffin, L. y Beck, AT. (2001) Evaluating geriatric medical outpatients with the Beck Depression Inventory Fast Screen for medical patients. *Aging Ment Health*, 5(2), 143-8.
- Steer, R., Cavalieri, D., Leonard, M., y Beck, A.T., (1999). Use of the Beck Depression Inventory for Primary Care to Screen for Major Depression Disorders. *General Hospital Psychiatry* 21, 106–111.
- Stegenga, B., Kamphuis, M., King, M., Nazareth, I. y Geerlings, M. (2010). The natural course and outcome of major depressive disorder in primary care: The PREDICT-NL study. *Social Psychiatry and Psychiatric Epidemiology Journal* (no se especifica página ni volumen). Disponible libre acceso en: www.springerlink.com.
- Stouffer, S.A., Guttman, L., Suchman, E.L., Lazarfeld, P.F., Star, S.A. y Clanssen J.S. (1950). *Measurement and prediction (Studies in Social Psychology in World War II)*. Princeton: Princeton University Press.
- Streiner, D. L., y Cairney, J. (2007). What's under the ROC? An introduction to Receiver Operating Characteristics curves. *The Canadian Journal of Psychiatry*, 52, 121-128.
- Summerfield, D. y Veale, D. (2008) Proposals for massive expansion of psychological therapies would be counterproductive across society. *The British Journal of Psychiatry*. 192, 326–330.
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285-1293.
- Tejero, A, Guimerá, E.M., Farré, J.M. y Peri J.M. (1986). Uso clínico del HAD (Hospital Anxiety and Depression Scale) en población psiquiátrica: un estudio de su sensibilidad, fiabilidad y validez. *Revista del Departamento de Psiquiatría de la Facultad de Medicina de Barcelona*, 13, 233-238.

- Terol, M.C., López-Roig, S., Rodríguez Marín, J., Martín-Aragón M., Pastor, M.A. y Reig M.T. Propiedades psicométricas de la Escala Hospitalaria de Ansiedad y Depresión (HAD). (2007). *Ansiedad y Estrés*, 13, (2-3), 163-176.
- Tomás, L. M., Meléndez, J. C., Oliver, A., Navarro, E., y Zaragoza, G. (2010). Efectos de método en las escalas de Ryff: un estudio en población de personas mayores. *Psicológica*, 31, 383-400.
- Tylee, A., Gastpar, M., Lepine, J.P. y Mendlewicz, J; (1999). DEPRES II (Depression Research in European Society II): a patient survey of the symptoms, disability and current management of depression in the community. DEPRES Steering Committee. *International Clinical Psychopharmacology*, 14,139-151.
- Urbina, J.R., Flores, J.M., García, M.P., Torres, L., y Torrubias, R.M. (2007)..
Síntomas depresivos en personas mayores. Prevalencia y factores asociados. *Gaceta Sanitaria*, 21(1):37-42.
- U.S. Preventive Services Task Force (2002). Screening for Depression:
Recommendations and Rationale. *Annals of Internal Medicine*, 136, 760-764.
- Van der Heiden, C., Muris, P., Bos, A. E. R., y Van der Molen, H. T. (2010). Factor structure of the Dutch version of the Penn State Worry Questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 41, 304-309.
- Vázquez, C., y Jiménez F. (2000). *Depresión y manía*. En Medición clínica en psiquiatría y psicología. Bulbena A., Berrios G., Fernández de Larrinoa P. Ed. (P.255-288). Barcelona: Masson.
- Vázquez, C., y Sanz, J. (1997). Fiabilidad y validez de la versión española del Inventario para la Depresión de Beck de 1978. *Clínica y Salud*, 7, 403-422.
- Vázquez, C., y Sanz, J. (1999). Fiabilidad y validez de la versión española del Inventario para la Depresión de Beck de 1978 en pacientes con trastornos psicológicos. *Clínica y Salud*, 10, 59-81.
- Vollebergh, WA.,Iedema, J.,Bijl, RV., de Graaf, R.,Smit, F.,Ormel, J.(2001).The structure and stability of common mental disorders: the NEMESIS study. *Archives of General Psychiatry*, 58, (6):597-603.
- Wetherell, J. L., y Arean, P. A. (1997). Psychometric evaluation of the Beck Anxiety Inventory with older medical patients. *Psychological Assessment*, 9, 136-144.

- Williams, J.B., Kobak, K.A., Bech, P., Engelhardt, N., Evans, K., Lipsitz, J., et al.
(2008) The GRID-HAMD: standardization of the Hamilton Depression Rating Scale. *International Clinical of Psychopharmacology*, 23, 120–9.
Disponible en <http://www.iscdd.org>
- Wilson, J.M.G. y Jungner, G. (1968). *Principles and practice*. World Health
No.34 Geneva: Organization Public Health Papers.
- World Health Organization (1992). *ICD-10 International classification of Mental
and Behavioural Disorders. Clinical Descriptions and diagnostic guidelines*.
Geneve: WHO.
- World Health Organization (2008). *Integrating mental health into primary care.
A global perspective*. (pp 1-224). Singapore: WHO and WONCA editors.
[http://www.who.int/mental_health/policy/Mental_health + primary care-20finalow-
res_120109.pdf](http://www.who.int/mental_health/policy/Mental_health+_primary_care-20finalow-res_120109.pdf)
- World Health Organization, (2007). Integrating mental health services into primary
health care. Geneva: Mental Health Policy, Planning and Service. Development
Information Sheet, Sheet 3.
Disponible en: http://www.who.int/mental_health/policy/services/en/index.htm
- Williams, J., Hitchcock, N., Cordes, J., Ramirez, G. y Pignone, M., (2002). Is this
patient clinically depressed?. *Journal of American Medicine Association*, 6,
(287), 9, 1160-1170.
- Wittkamp, K.A., Naeije, L., Schene, A.H., Huyser, J. y Van Weert, H.C.. (2007).
Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a
systematic review. *General Hospital Psychiatry*, 29, 388–95.
- Wilheim, K., Kotze, B., Kazdi-Pavlovick, D. y Parker, G. (2004). Screening for
Depression in the Medically Ill: A Comparison of Self-Report Measures,
Clinician Judgment, and DSM-IV Diagnoses. *Psychosomatics*; 45, 461–469.
- Winter, L., Steer, R., Hicks LJ., y Beck, A. (1999). Screening for Major Depression
Disorders in Adolescent Medical Outpatients With the Beck Depression
Inventory for Primary Care. *Journal of adolescent Health*, 24, 389–394.

- Wittchen, H.U., Holsboer, F. y Jacobi, F. (2001). Met and unmet needs in the management of depressive disorder in the community and Primary Care: The size and breadth of the problem. *Journal of Clinical Psychiatry*, 62, (26), 23-8.
- Yin, P. y Fan, X. (2000). Assessing the reliability of Beck Depression Inventory scores: Reliability generalization across studies. *Educational and Psychological Measurement*, 60, 201-223.
- Zung WWK. (1965) A self Rating depression scale. *Arch Gen Psychiatry*; 12. 63-70.
- Zigmond, A.S., y Snaith, R.P.(1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361–370.
- Zanarini, M. C., Skodol, A. E., Bender, D., Dolan, R., Sanislow, C., Schaefer, E., Morey, L. C., Grilo, C. M., Shea, M. T., McGlashan, T. H., y Gunderson, J. G. (2000). The Collaborative Longitudinal Personality Disorder Study: reliability of axis I and II diagnoses. *Journal of Personality Disorders*, 14, 291-299.
- Zastrow, A, Faude, V, Seyboth, F, Niehoff, D, Herzog, W, Löwe, B. (2008). Risk factors of symptom underestimation by physicians. *Journal of Psychosomatic Research.*, 64, 543–51.
- Zwick, W. R., y Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 3, 432-442.

Anexos

1.1. ANEXOS DEL ESTUDIO 1

1.1.1. Consentimiento informado

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

PROYECTO DE INVESTIGACIÓN SOBRE ALTERACIONES DEL ESTADO DE ÁNIMO EN PACIENTES DE ATENCIÓN PRIMARIA

DECLARACIÓN DE PARTICIPACIÓN VOLUNTARIA

Por este documento consiento en participar en este estudio sobre alteraciones en el estado de ánimo en pacientes de atención primaria.

Entiendo que realizaré una entrevista y varios cuestionarios donde se me harán preguntas sobre mis pensamientos, comportamientos y sentimientos. Entiendo que mis respuestas serán totalmente confidenciales y anónimas, y que puedo no contestar o dejar en blanco cualquier pregunta que no quisiera responder.

He sido avisado que el entrevistador responderá a cualquiera de mis preguntas sobre los procedimientos y propósito de este estudio, pero que ciertas preguntas que pueden afectar a los resultados de este estudio no me serán respondidas hasta después de la finalización del mismo.

He sido informado y comprendo que el estudio no pretende recoger información sobre individuos específicos, y que todos los datos serán codificados de forma que se mantendrá mi anonimato. Consiento en que los resultados de este estudio puedan ser publicados en revistas o libros científicos o difundidos por otros medios a la comunidad científica. No obstante, entiendo que mi nombre nunca aparecerá en dichos medios, que los informes de investigación sólo reflejarán los resultados del grupo y que la identidad de los participantes será protegida.

He leído y entendido las explicaciones mencionadas arriba y consiento en participar en este estudio.

Nombre del participante: _____

Fecha: _____

Firma del participante: _____

Certifico que he presentado esta información al participante y he obtenido su consentimiento.

Firma del investigador: _____

1.1.2. Protocolo evaluación

1.1.2.1. Datos sociodemográficos

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

DATOS SOCIODEMOGRÁFICOS

A. Nombre:

B. Edad:

C. Sexo (rodee con un círculo la opción que refleja su respuesta):

1. Varón
2. Mujer

D. Estado civil (rodee con un círculo la opción que refleja su respuesta):

1. Sotero/a
2. Casado/a o conviviendo de forma estable con una pareja
3. Separado/a o divorciado/a
4. Viudo/a

E. Estudios de más alto nivel oficial que Ud. haya cursado (con independencia de que los haya terminado o no) (rodee con un círculo la opción que refleja su respuesta):

1. No ha ido a la escuela o cursado ningún tipo de estudios
2. E.S.O, Estudios primarios, EGB o equivalente
3. Bachillerato, BUP, formación profesional o equivalentes
4. Estudios superiores universitarios o equivalentes

F. ¿En cuál de las siguientes situaciones laborales se encuentra Ud.? (rodee con un círculo la opción que refleja su respuesta):

1. Trabaja
2. Jubilado o pensionista
3. Parado
4. Estudiante
5. Sus labores
6. Otra situación

G. Profesión (rodee con un círculo la opción que refleja su respuesta):

1. Profesionales, técnicos o similares
2. Altos directivos de la Administración Pública o de la empresa privada
3. Propietarios o gerentes de la hostelería. Comercio y agricultura
4. Mandos intermedios, profesores y técnicos de la Administración Pública o de la empresa privada
5. Capataces y personal encargado en empresas o actividades de servicio
6. Personal administrativo y similar
7. Comerciantes, vendedores o similares
8. Personal de servicio
9. Trabajadores cualificados o semicualificados
10. Trabajadores agrícolas y no cualificados no agrícolas
11. Profesionales de las Fuerzas Armadas
12. Estudiante
13. Sus Labores
14. Otra ocupación

¿Está tomando en este momento alguna medicación?

1. NO
 2. Sí
- ¿Cuál?:.....
-
-
-

**1.1.2.2. Documento de diagnóstico del médico de atención
primaria**

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

PARA CUMPLIMENTAR POR EL MÉDICO DE ATENCIÓN PRIMARIA

Motivo de consulta:

.....

.....

.....

.....

Juicio diagnóstico actual:

.....

.....

.....

**Juicios diagnósticos de los últimos 6 meses: (especificar antecedentes
psiquiátricos):**

.....

.....

.....

.....

1.1.2.3. Inventario de Depresión de Beck–Cribado Rápido (BDI–FS)

Debido a restricciones de los derechos de autor del BDI-FS, sólo se reproducen algunos ítems del BDI-FS a modo de muestra. Para consultar íntegramente el BDI-FS o para su utilización, se puede contactar con *Pearson Educación, S. A.*, a través de su división *Pearson Clinical and Talent Assessment España* (<http://www.pearsonpsychcorp.es>).

BDI-FS para pacientes médicos

FECHA:

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

INSTRUCCIONES: Este cuestionario consiste en grupos de afirmaciones. Por favor, lea con atención cada uno de ellos y, a continuación, señale cuál de las afirmaciones de cada grupo describe mejor el modo en el que se ha sentido **DURANTE LAS DOS ÚLTIMAS SEMANAS, INCLUYENDO EL DÍA DE HOY**. Rodee con un círculo el número que se encuentre escrito a la izquierda de la afirmación que haya elegido. Si dentro del mismo grupo, hay más de una afirmación que considere igualmente aplicable a su caso, señálela también. **Asegúrese de leer todas las afirmaciones dentro de cada grupo antes de efectuar la elección.**

<p>1.</p> <p>0 No me siento triste habitualmente.</p> <p>1 Me siento triste gran parte del tiempo.</p> <p>2 Me siento triste continuamente.</p> <p>3 Me siento tan triste o tan desgraciado que no puedo soportarlo.</p>	<p>5.</p> <p>0 Siento lo mismo que antes sobre mí mismo.</p> <p>1 He perdido confianza en mí mismo.</p> <p>2 Estoy decepcionado conmigo mismo.</p> <p>3 No me gusta.</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

1.1.2.4. Escala Hospitalaria de Ansiedad y Depresión (HAD)

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

Escala Hospitalaria de Ansiedad y Depresión (HAD)

Los médicos conocen la importancia de los factores emocionales en la mayoría de enfermedades. Si el médico sabe cuál es el estado emocional del paciente puede prestarle entonces mejor ayuda. Este cuestionario ha sido confeccionado para ayudar a que su médico sepa cómo se siente usted afectiva y emocionalmente. No es preciso que preste atención a los números que aparecen a la izquierda. Lea cada pregunta y subraye la respuesta que usted considere que coincide con su propio estado emocional **en la última semana**. No es necesario que piense mucho tiempo cada respuesta; en este cuestionario las respuestas espontáneas tienen más valor que las que se piensan mucho.

A.1. Me siento tenso/a o nervioso/a:

- 3. Casi todo el día
- 2. Gran parte del día
- 1. De vez en cuando
- 0. Nunca

D.1. Sigo disfrutando de las cosas como siempre:

- 0. Ciertamente, igual que antes
- 1. No tanto como antes
- 2. Solamente un poco
- 3. Ya no disfruto con nada

A.2. Siento una especie de temor como si algo malo fuera a suceder:

- 3. Sí, y muy intenso
- 2. Sí, pero no muy intenso
- 1. Sí, pero no me preocupa
- 0. No siento nada de eso

D.2. Soy capaz de reírme y ver el lado gracioso de las cosas:

- 0. Igual que siempre
- 1. Actualmente, algo menos
- 2. Actualmente, mucho menos
- 3. Actualmente, en absoluto

A.3. Tengo la cabeza llena de preocupaciones:

- 3. Casi todo el día
- 2. Gran parte del día
- 1. De vez en cuando
- 0. Nunca

D.3. Me siento alegre:

- 3. Nunca
- 2. Muy pocas veces
- 1. En algunas ocasiones
- 0. Gran parte del día

Escala Hospitalaria de Ansiedad y Depresión (HAD) (Continuación)

A.4. Soy capaz de permanecer sentado/a tranquilo/a y relajado/a:

- 0. Siempre
- 1. A menudo
- 2. Raras veces
- 3. Nunca

D.4. Me siento lento/a y torpe:

- 3. Gran parte del día
- 2. A menudo
- 1. A veces
- 0. Nunca

A.5. Experimento una desagradable sensación de «nervios y hormigueos» en el estómago:

- 0. Nunca
- 1. Sólo en algunas ocasiones
- 2. A menudo
- 3. Muy a menudo

D.5. He perdido el interés por mi aspecto personal:

- 3. Completamente
- 2. No me cuido como debería hacerlo
- 1. Es posible que no me cuide como debiera
- 0. Me cuido como siempre lo he hecho

A.6. Me siento inquieto/a como si no pudiera parar de moverme:

- 3. Realmente mucho
- 2. Bastante
- 1. No mucho
- 0. En absoluto

D.6. Espero las cosas con ilusión:

- 0. Como siempre
- 1. Algo menos que antes
- 2. Mucho menos que antes
- 3. En absoluto

A.7. Experimento de repente sensaciones de gran angustia o temor:

- 3. Muy a menudo
- 2. Con cierta frecuencia
- 1. Raramente
- 0. Nunca

D.7. Soy capaz de disfrutar con un buen libro o con un buen programa de radio o televisión:

- 0. A menudo
- 1. Algunas veces
- 2. Pocas veces
- 3. Casi nunca.

ANEXOS DEL ESTUDIO 2

2.1. Consentimiento informado

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

PROYECTO DE INVESTIGACIÓN SOBRE ALTERACIONES DEL ESTADO DE ÁNIMO EN PACIENTES DE ATENCIÓN PRIMARIA

DECLARACIÓN DE PARTICIPACIÓN VOLUNTARIA

Por este documento consiento en participar en este estudio sobre alteraciones en el estado de ánimo en pacientes de atención primaria.

Entiendo que realizaré una entrevista y varios cuestionarios donde se me harán preguntas sobre mis pensamientos, comportamientos y sentimientos. Entiendo que mis respuestas serán totalmente confidenciales y anónimas, y que puedo no contestar o dejar en blanco cualquier pregunta que no quisiera responder.

He sido avisado que el entrevistador responderá a cualquiera de mis preguntas sobre los procedimientos y propósito de este estudio, pero que ciertas preguntas que pueden afectar a los resultados de este estudio no me serán respondidas hasta después de la finalización del mismo.

He sido informado y comprendo que el estudio no pretende recoger información sobre individuos específicos, y que todos los datos serán codificados de forma que se mantendrá mi anonimato. Consiento en que los resultados de este estudio puedan ser publicados en revistas o libros científicos o difundidos por otros medios a la comunidad científica. No obstante, entiendo que mi nombre nunca aparecerá en dichos medios, que los informes de investigación sólo reflejarán los resultados del grupo y que la identidad de los participantes será protegida.

He leído y entendido las explicaciones mencionadas arriba y consiento en participar en este estudio.

Nombre del participante: _____

Fecha: _____

Firma del participante: _____

Certifico que he presentado esta información al participante y he obtenido su consentimiento.

Firma del investigador: _____

2.2. Protocolo evaluación

2.2.1. Datos sociodemográficos

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

DATOS SOCIODEMOGRÁFICOS

A. Nombre:

B. Edad:

C. Sexo (rodee con un círculo la opción que refleja su respuesta):

1. Varón
2. Mujer

D. Estado civil (rodee con un círculo la opción que refleja su respuesta):

1. Sotero/a
2. Casado/a o conviviendo de forma estable con una pareja
3. Separado/a o divorciado/a
4. Viudo/a

E. Estudios de más alto nivel oficial que Ud. haya cursado (con independencia de que los haya terminado o no) (rodee con un círculo la opción que refleja su respuesta):

1. No ha ido a la escuela o cursado ningún tipo de estudios
2. E.S.O, Estudios primarios, EGB o equivalente
3. Bachillerato, BUP, formación profesional o equivalentes
4. Estudios superiores universitarios o equivalentes

F. ¿En cuál de las siguientes situaciones laborales se encuentra Ud.? (rodee con un círculo la opción que refleja su respuesta):

1. Trabajo
2. Jubilado o pensionista
3. Parado
4. Estudiante
5. Sus labores
6. Otra situación

G. Profesión (rodee con un círculo la opción que refleja su respuesta):

1. Profesionales, técnicos o similares
2. Altos directivos de la Administración Pública o de la empresa privada
3. Propietarios o gerentes de la hostelería. Comercio y agricultura
4. Mandos intermedios, profesores y técnicos de la Administración Pública o de la empresa privada
5. Capataces y personal encargado en empresas o actividades de servicio
6. Personal administrativo y similar
7. Comerciantes, vendedores o similares
8. Personal de servicio
9. Trabajadores cualificados o semicualificados
10. Trabajadores agrícolas y no cualificados no agrícolas
11. Profesionales de las Fuerzas Armadas
12. Estudiante
13. Sus Labores
14. Otra ocupación

¿Está tomando en este momento alguna medicación?

1. NO
2. SÍ

¿Cuál?:.....

.....
.....

2.2.2. Inventario de Depresión de Beck–Cribado Rápido (BDI–FS)

Debido a restricciones de los derechos de autor del BDI-FS, sólo se reproducen algunos ítems del BDI-FS a modo de muestra. Para consultar íntegramente el BDI-FS o para su utilización, hay que ponerse en contacto con *Pearson Educación, S. A.*, a través de su división *Pearson Clinical and Talent Assessment España* (<http://www.pearsonpsychcorp.es>).

BDI-FS para pacientes médicos

FECHA:

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

INSTRUCCIONES: Este cuestionario consiste en grupos de afirmaciones. Por favor, lea con atención cada uno de ellos y, a continuación, señale cuál de las afirmaciones de cada grupo describe mejor el modo en el que se ha sentido **DURANTE LAS DOS ÚLTIMAS SEMANAS, INCLUYENDO EL DÍA DE HOY**. Rodee con un círculo el número que se encuentre escrito a la izquierda de la afirmación que haya elegido. Si dentro del mismo grupo, hay más de una afirmación que considere igualmente aplicable a su caso, señálela también. **Asegúrese de leer todas las afirmaciones dentro de cada grupo antes de efectuar la elección.**

1.

- 0 No me siento triste habitualmente.
- 1 Me siento triste gran parte del tiempo.
- 2 Me siento triste continuamente.
- 3 Me siento tan triste o tan desgraciado que no puedo soportarlo.

5.

- 0 Siento lo mismo que antes sobre mí mismo.
- 1 He perdido confianza en mí mismo.
- 2 Estoy decepcionado conmigo mismo.
- 3 No me gusta.

2.2.3. Escala Autoaplicada para la Depresión de Zung (SDS)

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

ESCALA AUTOAPLICADA PARA LA DEPRESIÓN DE ZUNG Y CONDE

Instrucciones: A continuación se expresan varias respuestas posibles a cada uno de los 20 apartados. Detrás de cada frase marque con una cruz la casilla que mejor refleje su situación actual.

- A = Muy poco tiempo/ Muy pocas veces/ Raramente.
 B = Algún tiempo/ Algunas veces/ De vez en cuando.
 C = Gran parte del tiempo/ Muchas veces/ Frecuentemente.
 D = Casi siempre/ Siempre/ Casi todo el tiempo.

Me siento triste y deprimido/a.	A	B	C	D
Por las mañanas me siento mejor que por las tardes.				
Frecuentemente tengo ganas de llorar y a veces lloro.				
Me cuesta mucho dormir o duermo mal por la noche.				
Ahora tengo tanto apetito como antes.				
Todavía me siento atraído/a por el sexo opuesto.				
Creo que estoy adelgazando.				
Estoy estreñado/a				
Tengo palpitaciones.				
Me canso por cualquier cosa.				
Mi cabeza está tan despejada como antes.				
Hago las cosas con la misma facilidad que antes.				
Me siento agitado/a e intranquilo/a y no puedo estar quieto/a.				
Tengo esperanza y confianza en el futuro.				
Me siento más irritable que habitualmente.				
Encuentro fácil tomar decisiones.				
Me creo útil y necesario para la gente.				
Encuentro agradable vivir, mi vida es plena.				
Creo que sería mejor para los demás que me muriera.				
Me gustan las mismas cosas que habitualmente me agradaban.				

2.2.4. Inventario de Ansiedad de Beck (BAI)

Debido a restricciones de los derechos de autor del BAI, sólo se reproducen algunos ítems del BAI a modo de muestra. Para consultar íntegramente el BAI o para su utilización, se puede contactar con *Pearson Educación, S. A.*, a través de su división *Pearson Clinical and Talent Assessment España* (<http://www.pearsonpsychcorp.es>).

CÓDIGO MAP/PARTICIPANTE

--	--	--	--	--

INVENTARIO DE ANSIEDAD DE BECK (BAI)

Instrucciones: A continuación hay una lista de síntomas frecuentes de ansiedad. Léalos atentamente e indique el grado en que se ha visto afectado por cada uno de ellos durante la última semana, rodeando el número que mejor lo indique en relación con la siguiente escala:

<p>0 = Nada en absoluto 1 = Levemente, no me molestó mucho. 2 = Moderadamente, fue muy desagradable pero podía soportarlo. 3 = Gravemente, casi no podía soportarlo.</p>

1. Hormigueo o entumecimiento.	0	1	2	3
2. Sensación de calor.	0	1	2	3
3. Debilidad en las piernas.	0	1	2	3
4. Incapacidad para relajarme.	0	1	2	3